

Towards Accelerated Transient Solvers for Full System Power Integrity Verification

Original

Towards Accelerated Transient Solvers for Full System Power Integrity Verification / Carlucci, Antonio; Grivet-Talocia, Stefano; Mongrain, Scott; Kulasekaran, Sid; Radhakrishnan, Kaladhar. - ELETTRONICO. - (2022), pp. 1-3. (2022 IEEE 31st Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS) San Jose, CA, USA 09-12 October 2022) [10.1109/EPEPS53828.2022.9947106].

Availability:

This version is available at: 11583/2974436 since: 2023-01-09T14:12:48Z

Publisher:

IEEE

Published

DOI:10.1109/EPEPS53828.2022.9947106

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

IEEE postprint/Author's Accepted Manuscript

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

Towards Accelerated Transient Solvers for Full System Power Integrity Verification

Antonio Carlucci*, Stefano Grivet-Talocia*, Scott Mongrain[§], Sid Kulasekaran[§], Kaladhar Radhakrishnan[§]

*Dept. Electronics and Telecommunications, Politecnico di Torino, Italy

[§]Intel Corporation, Chandler, AZ, USA

antonio.carlucci@polito.it

Abstract—This paper proposes a novel framework for power integrity verification of multicore systems, including voltage stabilization provided by multiple integrated voltage regulators at the core interfaces. The proposed framework adopts a two-stage macromodeling strategy to derive a compact representation of the full system dynamics as observed from each core. These dynamics are parameterized by the time-varying duty cycle provided by dedicated feedback controllers to each voltage regulator, here implemented through an averaged model. We show that the proposed simulation framework has the potential to outperform direct transient analysis based on SPICE engines.

I. INTRODUCTION AND PROBLEM STATEMENT

This paper addresses system level power integrity verification for multi-core microprocessors with integrated voltage regulators providing per-core voltage domain granularity. Our objective is to simulate an entire power delivery network structured as a cascade of multiple stages between the main supply and the compute domains inside the microprocessor [1]. In particular, we address the transient solution of the complete power delivery, including voltage regulation effects provided by Fully Integrated Voltage Regulators (FIVR) [2] as well as the coupling from one core to another due to the shared input network. The large-scale nature of this simulation problem, both in terms of expected dynamic order and number of ports/signals to be evaluated, combined with the nonlinear FIVR circuitry and the associated feedback regulation loops, make this problem particularly challenging. A brute-force SPICE transient simulation based on suitable models for all system parts is feasible only for small-scale low-core platforms, but it does not have potential to scale to higher complexities, as demanded by state of the art HPC or AI many-core processors.

The reference structure is schematically depicted in Fig. 1. At the motherboard level, the output of an on-board Voltage Regulator Module (VRM) is routed through the PCB power planes and the power pins of the microprocessor package. The entire board and package power distribution structures are collectively denoted here as *input network*, which also includes a linear model of the VRM and the contribution of all board and package decoupling capacitors. The latter are usually optimized to meet specific target impedance requirements [3], [4], but once this optimization is performed they can be considered as integral parts of the input network, which in turn can be collectively represented a large-scale distributed Linear

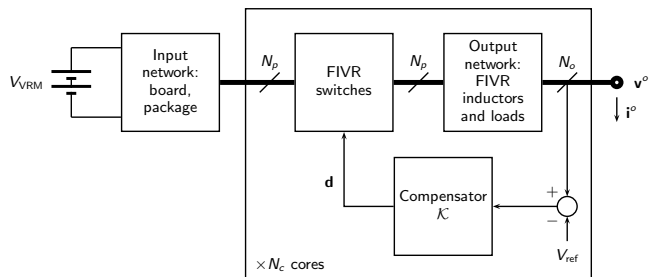


Fig. 1. Schematic illustration of the power distribution system under investigation, including N_c cores whose voltage is regulated by N_p -phase FIVRs.

and Time-Invariant (LTI) multiport described by a transfer matrix $\mathbf{H}_b(s)$.

Inside the chip, a second voltage regulation stage is implemented through FIVRs, consisting of multi-phase switching power supplies (e.g., buck converters). Power transistors, switching control circuits, and the output decoupling for these FIVRs are fabricated on-die, while the inductors are placed in the package. As a result, the FIVR output is a filtered and regulated voltage that is distributed through the die power rails to reach logic devices in their respective power domains. In this work, the switching banks of all FIVRs are represented by nonlinear and time-varying circuit blocks. The FIVR switches are controlled through feedback loops that sense the output voltage of each core and translate the tracking error with respect to a reference voltage V_{ref} into the appropriate duty cycle. This operation is attained through dedicated per-core controllers, denoted as \mathcal{K} in Fig. 1. The figure also shows blocks denoted as *output network* which include a circuit model of the PDN of each core, including integrated MIM capacitors which provide the output decoupling, plus a detailed electromagnetic model of the integrated inductors that complete the topology of the FIVRs. This output network can be represented as a LTI system with a transfer matrix $\mathbf{H}_c(s)$.

In order to set notation, we consider a system with N_c identical cores, whereas the number of phases of each FIVR is N_p . As a result, the input network $\mathbf{H}_b(s)$ has $N_c N_p + 1$ ports, each core k is represented by a transfer function $\mathbf{H}_{c,k}(s)$ with N_p ports interfaced to the switches and N_o output ports where the transient voltage is to be computed, so that the

overall output network $\mathbf{H}_c(s)$ has a total of $N_c(N_p + N_o)$ ports. The time-varying duty cycles of all cores are collected in the vector $\mathbf{d}(t) \in [0, 1]^{N_c}$. The main objective of this work is to compute efficiently the transient voltages $v_{k,n}^o(t)$ at all $n = 0, \dots, N_o - 1$ ports of each core $k = 0, \dots, N_c - 1$, excited by predefined current load signals $i_{k,n}^o(t)$ acting concurrently. All results in this work refer to a model of a 11th Generation Intel® Core™ microprocessor, for which $N_c = 4$, $N_p = 4$ and $N_o = 36$. The specific models that are used in this work include only partial output decoupling, as will be evident from AC and transient results.

II. FORMULATION

The initial step in our problem setup is to properly characterize both the input network and the output network. Since these are viewed as LTI subsystems, we follow the standard practice of performing a set of full-wave electromagnetic analyses of the distributed interconnects (board+package) and components (e.g. the FIVR inductor banks), obtaining sampled scattering responses at the ports of interest. These are combined with any lumped terminations (e.g. decoupling capacitors), and the resulting assembled scattering samples are processed by a rational macromodeling engine based on Vector Fitting (VF) with passivity enforcement [5], [6], so that both $\mathbf{H}_b(s)$ and $\mathbf{H}_{c,k}(s)$ are available as a set of linear state-space equations and the associated synthesized SPICE realization. Such macromodels enable a direct (reference) SPICE simulation of the complete system, once complemented with circuit models of the switches and the compensators. This will provide the solution that we will use as reference, both in terms of accuracy and runtime.

One of the key aspects of proposed formulation is the adopted representation for the switches, here represented through averaged models. For each core k and phase j , the corresponding set of FIVR switches is represented by an ideal transformer with turn ratio $1 : d_k(t)$, where $d_k(t)$ is the duty cycle signal resulting from the compensator \mathcal{K}_k of core k . This assumption has its own limitations but is known to be accurate when the buck converters operate in Continuous Conduction Mode (CCM). In particular, if each duty cycle signal $d_k(t)$ is “frozen” and considered as a fixed parameter (e.g. by disconnecting the controllers and opening the feedback loops), the entire structure becomes a large-scale LTI system, which can be fully characterized by the output impedance matrix $\mathbf{Z}(s, \mathbf{d})$ relating the output voltages to the output excitation currents through $\mathbf{V}^o(s) = \mathbf{Z}(s, \mathbf{d})\mathbf{I}^o(s)$.

Figure 2 depicts a sweep over frequency and duty cycle of two representative output impedance elements. The particular structure of such responses and their dependence on the duty cycles \mathbf{d} enable a second layer of model order reduction through a second rational fitting stage, where common poles p_ν are used to represent all impedance entries, and where the associate residues \mathbf{R}_ν are parameterized

$$\mathbf{Z}(s, \mathbf{d}) = \sum_{\nu=1}^{\bar{\nu}} \frac{\mathbf{R}_\nu(\mathbf{d})}{s - p_\nu}. \quad (1)$$

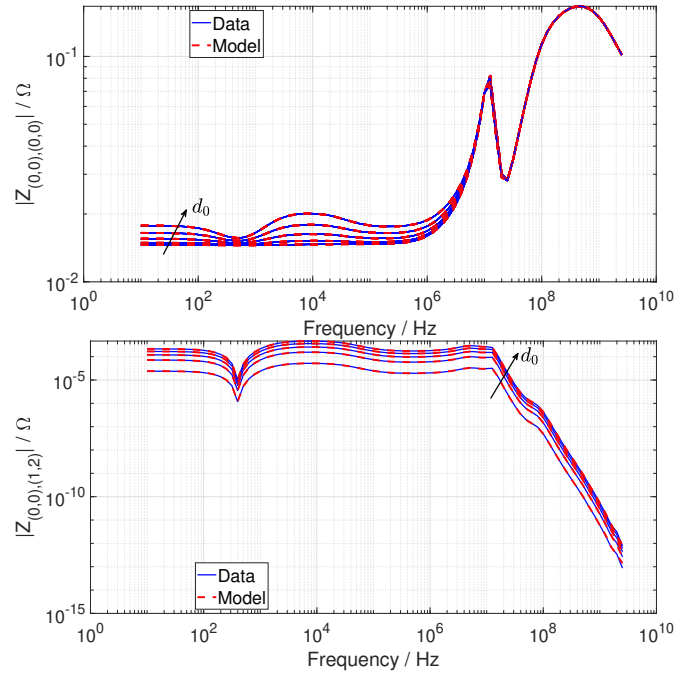


Fig. 2. Open-loop output impedance responses $Z^{(k_1, \ell_1)(k_2, \ell_2)}(j\omega, \mathbf{d})$, at port ℓ_1 of core k_1 while exciting port ℓ_2 of core k_2 , plotted over a sweep of duty cycle values d_0 at fixed $d_1 = d_2 = d_3 = 0.1$. The macromodel responses (dashed lines) are compared to reference AC sweeps from HSPICE (thin solid lines).

A closed-form parameterization of the residues is obtained by a low-order polynomial interpolation. A state-space realization of the impedance (1), including also the contribution of the constant input V_{VRM} , can be obtained as

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_o(\mathbf{d})\mathbf{i}^o(t) + \mathbf{B}_i(\mathbf{d})V_{\text{VRM}} \\ \mathbf{v}^o(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \quad (2)$$

where the dependence on the duty cycle is not affecting the open-loop dynamics (matrix \mathbf{A} collecting the poles p_ν is constant) but only the input-state mappings $\mathbf{B}_{i,o}$.

The final step in proposed framework is to reintroduce the closed loop control of the duty cycle signals. This is achieved through the following system of ODEs

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_o(\mathbf{d}(t))\mathbf{i}^o(t) + \mathbf{B}_i(\mathbf{d}(t))V_{\text{VRM}} \\ \mathbf{v}^o(t) = \mathbf{C}\mathbf{x}(t) \\ \dot{\mathbf{w}}(t) = \mathbf{A}_\mathcal{K}\mathbf{w}(t) + \mathbf{B}_\mathcal{K}\mathbf{e}(t) \\ \mathbf{d}(t) = \mathbf{C}_\mathcal{K}\mathbf{w}(t) + \mathbf{D}_\mathcal{K}\mathbf{e}(t) \end{cases} \quad (3)$$

where the dynamics of all compensators are represented as a (vectorized) state-space system (subscript \mathcal{K}). The vector of error signals $\mathbf{e}(t) \in \mathbb{R}^{N_c}$ feeding the compensators is defined as $\mathbf{e}(t) = \mathbf{N}\mathbf{v}^o(t) - \mathbf{V}_{\text{ref}}$, where \mathbf{N} is a constant selector matrix and \mathbf{V}_{ref} collects the reference voltages. Note that the explicit time-dependence of all signals is highlighted in (3), so that it becomes evident that the system is not represented as a standard LTI but rather as a nonlinear system in Linear Parameter Varying (LPV) form with feedback. The nonlinearity is however only algebraic (polynomial).

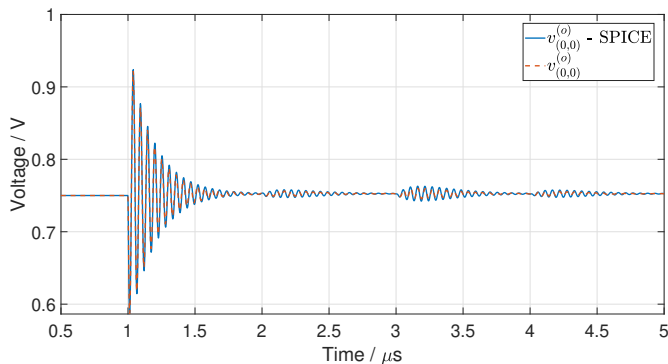


Fig. 3. Transient response of the regulated voltage $v_{k,\ell}^o(t)$ at core $k = 0$, port $\ell = 0$ induced by a sequential transient current step (10 A / 5 ns) excitation per core. The proposed solver response (dashed line) is compared to the reference HSPICE response (thin solid line).

In order to integrate the above system of ODEs numerically, we consider a uniform time step Δt with $t_n = n\Delta t$ for $n = 0, 1, \dots, N_t$, and we initialize all system states at t_0 with their DC solution (no output current excitation and all voltages equal to the V_{ref}). Denoting the approximation induced by discretization as $\hat{\mathbf{x}}_n \approx \mathbf{x}(t_n)$, we introduce an additional relaxation assuming that $\mathbf{d}(t)$ and $\mathbf{e}(t)$ are piecewise constant in each sub-interval, that is $\hat{\mathbf{d}}_n \approx \mathbf{d}(t_n) \approx \mathbf{d}(t)$ for all $t \in [t_n, t_{n+1}]$, and similarly for $\mathbf{e}(t)$. This approximation enables the application of recursive convolutions to integrate equations (3) from t_n to t_{n+1} , where the ODEs are locally linear and the PDN system is decoupled from \mathcal{K} . Differently from standard rational macromodel transient simulation approaches [5], the coefficients of each recursive convolution are time-dependent and updated at each time step. In a compact form, the proposed discretized solution reads

$$\begin{aligned} \hat{\mathbf{x}}_{n+1} &= e^{\mathbf{A}\Delta t} \hat{\mathbf{x}}_n + \\ &\int_{t_n}^{t_{n+1}} e^{\mathbf{A}(t_{n+1}-\tau)} \left[\mathbf{B}_o(\hat{\mathbf{d}}_n) \mathbf{i}^o(\tau) + \mathbf{B}_i(\hat{\mathbf{d}}_n) V_{\text{VRM}} \right] d\tau \\ \hat{\mathbf{v}}_n^o &= \mathbf{C} \hat{\mathbf{x}}_n \\ \hat{\mathbf{w}}_{n+1} &= e^{\mathbf{A}_{\mathcal{K}} \Delta t} \hat{\mathbf{w}}_n + \int_{t_k}^{t_{n+1}} e^{\mathbf{A}_{\mathcal{K}}(t_{n+1}-\tau)} \mathbf{B}_{\mathcal{K}} \hat{\mathbf{e}}_n d\tau \\ \hat{\mathbf{d}}_n &= \mathbf{C}_{\mathcal{K}} \hat{\mathbf{w}}_n + \mathbf{D}_{\mathcal{K}} \hat{\mathbf{e}}_n \end{aligned}$$

III. NUMERICAL EXPERIMENTS

The scheme proposed here has proven effective for transient simulation of the power delivery network of a 4-core microprocessor. In this real-world test case, we are able to show that a prototypal, non-optimized and non-parallel Matlab implementation of this solver is already about $10\times$ faster than commercial circuit solvers like HSPICE, while still providing accurate results for the purposes of power integrity verification.

Each of the four cores in the test case is a FIVR domain with $N_o = 36$ ports on the die side, totalling $P = 144$

output ports. The obtained reduced-order PDN model (1) has dynamic order $\bar{\nu} = 24$ and is parameterized in terms of the duty cycle \mathbf{d} with polynomial degrees $\rho_o = 2$ and $\rho_i = 1$ for the corresponding two terms in (2). The raw data used to build this macromodel are the PDN Z-parameters sampled for 625 values of \mathbf{d} arranged on a uniform grid in the parameter space, resulting from parametrically-swept AC analyses performed in HSPICE. The macromodel accuracy is demonstrated in Fig. 2.

The PDN system, initially at steady-state, is excited with a 10 A step per core with rise time 5 ns, uniformly distributed among all ports of each core. The individual cores are activated sequentially at $\{1, 2, 3, 4\} \mu\text{s}$, and a transient simulation is performed up to $T = 5 \mu\text{s}$ with a fixed time step $\Delta t = 0.1 \text{ ns}$. In order to perform a fair comparison, the total number of time steps is the same as in the reference HSPICE simulation. The transient results at one representative output port of proposed solver are reported in Fig. 3 where also the reference HSPICE solution is depicted for comparison. The RMS error in the output voltage with respect to the HSPICE transient simulation is 3.9 mV, corresponding to a relative 0.5% cumulative RMS error. In terms of runtime, proposed solver completed the transient analysis in 209 s whereas HSPICE required 1920 s, with a corresponding speedup factor of about $9.6\times$.

IV. CONCLUSIONS

This paper provided a proof of concept of a macromodel-based transient solver for full-system power integrity verification, including core voltage stabilization through averaged models of integrated regulators. The proposed approach has been validated on a model of the power distribution network of an Intel-based 4-core microprocessor, showing excellent accuracy with respect to reference SPICE simulation and confirming a good potential for a dramatic speedup. Future developments will be dedicated to code optimization and parallelization, as well as scalability to higher core counts.

REFERENCES

- [1] K. Radhakrishnan, M. Swaminathan, and B. K. Bhattacharyya, "Power delivery for high-performance microprocessors—challenges, solutions, and future trends," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 11, no. 4, pp. 655–671, 2021.
- [2] E. A. Burton, G. Schrom, F. Paillet, J. Douglas, W. J. Lambert, K. Radhakrishnan, and M. J. Hill, "FIVR — Fully integrated voltage regulators on 4th generation Intel® Core™ SoCs," in *2014 IEEE Applied Power Electronics Conference and Exposition - APEC 2014*, Mar. 2014, pp. 432–439.
- [3] M. Swaminathan and A. E. Engin, *Power integrity modeling and design for semiconductors and systems*. Prentice Hall Press, 2007.
- [4] I. Erdin and R. Achar, "Mcb-dpo: Multiport constrained barrier method-based decoupling capacitor placement optimization on irregularly shaped planes," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 12, no. 4, pp. 665–675, 2022.
- [5] S. Grivet-Talocia and B. Gustavsen, *Passive macromodeling: Theory and applications*. John Wiley & Sons, 2015.
- [6] "IdEM R2018, Dassault Systèmes." [Online]. Available: www.3ds.com/products-services/simulia/products/idem/