

**SMART Grid On Chip: Infusing intelligence to
on-chip energy management**

A Thesis

Submitted to the Faculty

of

Drexel University

by

Divya Pathak

in partial fulfillment of the

requirements for the degree

of

Doctor of Philosophy

November 2018

Supervised by

Professor Ioannis Savidis





GRADUATE THESIS/DISSERTATION APPROVAL FORM AND SIGNATURE PAGE

Instructions: This form must be completed by all master's and doctoral students with a thesis or dissertation requirement. Please type or print clearly as this form MUST be included as page 1 of your thesis or dissertation via electronic submission to ProQuest. All theses and dissertations must be formatted according to the University and department/program requirements. **Reminder:** It is the responsibility of the student to submit any/all edits requested by the Examining Committee to the Faculty Mentor or Supervising Professor for final approval and signature via the Graduate Program Completion Form.

Type:

☐

Master's Thesis

☒

PhD/Doctoral Thesis or Dissertation

Thesis or Dissertation Title: SMART GRID On Chip: Infusing intelligence to on-chip power management

Author's Name:

Divya Pathak

Month and Year:

November 2018

The signatures below certify that this thesis / dissertation (circle one) is complete and approved by the Examining Committee.

Committee Chairperson's Name:

Baris Taskin

Title: Professor

Department: ECE

Institution (if other than Drexel University):

Signature: B. Taskin

Committee Member's Name: Nagarajan Kandasamy

Title: Professor

Department: ECE

Institution (if other than Drexel University):

Signature: K. N. K.

Committee Member's Name: Antoniós Kotsos

Title: Associate Professor

Department: MEM

Institution (if other than Drexel University):

Signature: [Signature]

Committee Member's Name: Ioannis Savidis

Title: Assistant Professor

Department: ECE

Institution (if other than Drexel University):

Signature: [Signature]

Committee Member's Name: Anup Das

Title: Assistant Professor

Department: ECE

Institution (if other than Drexel University):

Signature: [Signature]

Committee Member's Name: _____

Title: _____

Department: _____

Institution (if other than Drexel University):

Signature: _____

© Copyright 2018
Divya Pathak. All Rights Reserved.

Dedications

My mother, Shammi Pathak for her unconditional love and patience.

*My father, Wg. Cdr. S. C. Pathak, for introducing me to engineering and making
me a knowledge seeker for life.*

Acknowledgments

“Energy and persistence can capture anything.” Having lived in the city of Philadelphia for the past five years, I feel obligated to start with a quote by Benjamin Franklin. So here I go expressing my gratitude to the people who have fueled my energy and supported my persistence in the quest to seek knowledge, which is now culminating into a Ph.D. in Computer Engineering.

I have to begin with my advisor Professor Ioannis Savidis. Our journey at Drexel University started together in the fall of 2013. I am truly honored to be his first Ph.D. student and cannot thank him enough for his support and encouragement in the last five years. I started my graduate studies after a long and successful stint in the semiconductor industry. Making a transition to a student, while being far from my family was not easy. I truly appreciate Professor Savidis’s faith in my scholastic abilities and engaging me in research early on while I was immersed with the workload of graduate courses. Professor Savidis, I have learned a lot from you about how to be a successful academician, how to run a research lab, how to write research grants, and the list is endless. Forever shall I remain indebted to you.

My sincere thanks to Professor Baris Taskin for chairing my Ph.D. dissertation advisory committee, for the short but insightful conversations and imparting me knowledge through his courses on electronic design automation. I truly admire his wit and knack to grasp research problems. I thank Professor Nagarajan Kandasamy, Professor

Anup Das and Professor Antonios Kontsos for serving on my Ph.D. dissertation advisory committee and providing me valuable feedback on my research. Your feedback has been very helpful in improving my dissertation.

I gained valuable research experience through our research collaboration with Professor Houman Homayoun at George Mason University. The research meetings with Professor Homayoun were always engaging and enlightening. I thank him and his graduate students for our successful research collaboration.

I would like to thank Professor Mark Hempstead for chairing my Ph.D. candidacy exam and teaching me the nuts and bolts of computer architecture through his course offering at Drexel University. The structure of the classes with the engaging research paper discussions and projects was mentally invigorating.

I spent two summers at IBM Thomas J. Watson research center as a research intern in the VLSI group. I feel blessed to have been granted the opportunity to work amongst giants in the field of circuit design and computer architecture. It was an honor to be in the same conference room as Dr. Robert H. Dennard. My sincere thanks to Dr. Phillip Restle for being my mentor and appreciating my research effort. Dr. Restle, I learned a lot from you. I sincerely hope to maintain the same zeal and enthusiasm towards research as you do twenty years from now. I thank all the members of the VLSI group for their guidance and support.

Special thanks to all my former and current ICE lab mates. Sharing lighter moments with you on various occasions made the journey towards my Ph.D. that much more easier.

I cannot thank enough the IEEE CAS Society for awarding me the 2017 IEEE

Circuits and Systems Pre-Doctoral Scholarship. Apart from the monetary award, the honor of being recognized by the CASS members means a lot to me. I am indebted to the generous benefactors of the Joseph and Shirley Carleone Endowed Fellowship, Allen Rothwarf Scholarship, and Seaman Fellowship for bestowing me with fellowships to aid my graduate studies at Drexel.

I would also like to acknowledge my former managers at STMicroelectronics, for honoring my decision to pursue graduate studies in the USA and providing me all the support to relieve me from my technical and administrative duties as a team leader. The strong problem solving skills that I developed during my various silicon bring up and debugging sessions at Noida, Grenoble, and Crolles have helped me tremendously as a researcher.

The acknowledgment would be incomplete without thanking Mr. Thomas Wagnagiris, my former IEEE mentor, who advised me to apply to Drexel University for graduate studies. The successful journey at Drexel would not have started without your guidance Tom. Thank you.

In the end, I cannot thank enough my family for bearing my absence during all moments of joy and tough times in the past five years. Your unwavering love and support has made me what I am today.

Table of Contents

LIST OF TABLES	xii
LIST OF FIGURES	xiv
ABSTRACT	xxii
1. INTRODUCTION	1
1.1 SMART GRID based on-chip power management	4
1.1.1 Supply side management	5
1.1.2 Demand side management	7
1.2 Outline of the dissertation	7
1.2.1 Research objective	8
1.2.2 Research methods	8
1.2.3 Research contributions	8
2. BACKGROUND ON DC-DC VOLTAGE CONVERTERS	12
2.1 On-chip DC-DC voltage regulator topologies	13
2.1.1 Switched capacitor voltage converters	15
2.1.2 Switching DC-DC voltage converters	16
2.1.3 Low drop out voltage regulators	18
2.1.4 Hybrid voltage regulators	21
2.2 Comparison between different OCVR topologies	21
2.3 Challenges in implementing on-chip voltage regulators	23

2.3.1	Improvements in on-chip voltage regulator design	23
2.3.2	Distributed on-chip voltage regulation	26
2.3.3	Adaptive voltage scaling	26
2.4	Summary	27
3.	POWER INTEGRITY IN SUB-20 NM TECHNOLOGIES	29
3.1	Nanometer scale field effect transistor technology	29
3.1.1	Process variation	31
3.1.2	MOS transistor aging	32
3.1.3	Transient faults	37
3.1.4	Temperature effect inversion and self heating in FinFETS	38
3.2	Electromigration	41
3.3	Combined effect of PVT, aging, SHE, TEI, and noise	41
3.4	Summary	45
4.	ON-CHIP POWER MANAGEMENT	46
4.1	Technology level power management	47
4.2	Circuit level power management	48
4.3	Architecture level power management	49
4.4	System level power management	51
4.5	Cross-layer power management	52
4.5.1	Cross layer adaptive voltage scaling	53
4.5.2	Cross layer error-resilient voltage over scaling	55
4.5.3	Thread to core mapping accounting for on-chip voltage regulator ef- ficiency	55

4.5.4	Core scaling with enhanced power management	57
4.6	Summary	59
5.	UNDER-PROVISIONED AND RECONFIGURABLE POWER DELIVERY	60
5.1	Related work	62
5.2	Power dissipation in CMPs	65
5.2.1	Simulation methodology	68
5.2.2	Analysis of the power dissipation behavior	68
5.3	Proposed interconnected power delivery network	70
5.4	Improvement in figures of merit of OCVRs and other system parameters	76
5.4.1	Improvement in figures of merit of a buck converter	79
5.4.2	Improvement in power conversion efficiency of a buck converter . . .	80
5.5	Energy efficiency of a CMP with RPDN and under-provisioned voltage regulators	84
5.6	Transient analysis of the RPDN	88
5.7	Summary	95
6.	WORK LOAD SCHEDULING FOR CHIP MULTIPROCESSORS WITH UNDER- PROVISIONED POWER DELIVERY	96
6.1	System model and notations	99
6.1.1	CMP models	99
6.1.2	Power model	100
6.1.3	Voltage regulator models	101
6.1.4	Real-time periodic task model	103
6.2	Optimal workload scheduling	103
6.3	Workload scheduling heuristic	104

6.4	Simulation results	106
6.5	Summary	111
7.	HYPERABRUPT VARACTORS FOR POWER SUPPLY NOISE MITIGATION . .	112
7.1	Clock-edge induced power supply noise	114
7.2	Hyperabrupt junction varactor diode	119
7.2.1	Modeling of a hyperabrupt junction varactor	121
7.2.2	Power supply noise suppression using hyperabrupt junction diodes .	125
7.3	SPICE simulations characterizing noise suppression	130
7.3.1	Construction of the power distribution network	131
7.3.2	Transient current waveforms for analysis of power supply noise . . .	134
7.4	Comparison with existing power supply noise mitigation techniques . .	138
7.5	Summary	142
8.	RUN-TIME POWER SUPPLY VOLTAGE DETECTION AND CLAMPING MECH- ANISM FOR 3-D ICs	144
8.1	Power supply voltage detection and clamping technique	146
8.2	Ring oscillator circuit for each voltage domain	147
8.3	Method I for power plane construction: Frequency to voltage converter topology	152
8.3.1	Frequency to voltage converter circuit	153
8.3.2	Implementation of the voltage controlled voltage source	154
8.3.3	Simulation results	155
8.4	Method II for power plane construction: Frequency comparator topology	157
8.4.1	Divide by 40 clock divider	158
8.4.2	Frequency comparator	158

8.4.3	Voltage ramp generator	159
8.4.4	Voltage peak detector	160
8.4.5	Simulated Circuit Results	161
8.5	Summary	163
9.	EVOLVING ON-CHIP POWER DELIVERY THROUGH PARTICLE SWARM OPTIMIZATION	165
9.1	Learning algorithm for an evolvable PDN	169
9.1.1	Voltage guard-band modulation based on particle swarm optimization	171
9.1.2	Problem formulation with PSO	172
9.1.3	Evolving voltage assignment	175
9.1.4	Timing and aging sensor data to direct the PSO	177
9.2	Simulated Results of PSO voltage assignment	182
9.2.1	Simulation setup	184
9.2.2	Voltage assignment by the PSO algorithm	187
9.2.3	Power supply noise compensation	190
9.2.4	Reduction in transistor aging	191
9.2.5	Reduction in operating temperature	195
9.2.6	Hardware overhead of run-time PSO implementation	197
9.3	Summary	198
10.	CONCLUSIONS	200
11.	FUTURE DIRECTIONS	204
11.1	Energy efficient domain specific architectures	204
11.2	Power management of swarm of IoT sensors for military applications	206

11.3 Applications to hardware security	208
BIBLIOGRAPHY	209
VITA	226

List of Tables

2.1	Comparison between few on-chip voltage regulator designs.	21
5.1	Architectural parameters of the core.	66
5.2	Summary of important system, architecture and circuit parameters used to design the RPDNs.	67
5.3	Combined power dissipation characteristics of SPEC CPU2000 and SPEC CPU2006 benchmarks.	69
5.4	Simulation parameters for <i>Algorithm 5.3a</i>	78
5.5	Improvement in the figures of merit of the OCVR supporting I_{avg} instead of I_{peak}	79
5.6	Power consumption of DC-DC switching buck converters designed for 0.6 A and 6 A peak load currents.	83
5.7	Properties of the switching DC-DC buck converters simulated for maximum power conversion efficiency.	83
5.8	Parameters of the PDN determined through SPICE simulation.	86
5.9	IBM power grid benchmark for transient analysis (<i>ibmpg1t</i>). The benchmark includes the number of current sources I , nodes N , resistors R , inductors L , capacitors C , shorts S (zero value resistors and voltage sources), and connections V to the 1.8 V off-chip voltage source. The V off-chip connections are replaced by a Verilog-A model of the OCVR with V connections to the on-chip PDN.	90
6.1	Frequency (MHz) and voltage (V) pairs used by the DFVS procedure in <i>Algorithm 6.3a</i> . The bold values listed in the table are the nominal frequencies and voltages.	99
6.2	Parameters of the CMP cores derived from the Samsung Exynos 5410 big.LITTLE architecture.	100
6.3	Operating values of the switching DC-DC buck converters serving the ARM A15 and A7 cores.	102

6.4	Parameters to generate real time periodic tasksets.	106
7.1	Parameters used to simulate the switched capacitor load circuit.	117
7.2	SPICE model parameters for the hyperabrupt varactor diode SMVA-001LF.	124
7.3	Components of the optimized PDN.	133
7.4	Characterization of leakage current, capacitor density, and temperature stability for different families of capacitors.	142
9.1	Results stored in a look up table (LUT_i) for the i^{th} timing sensor.	185
9.2	Architectural parameters of the core in a 16 core CMP.	187
9.3	Technology and power-law fitting parameters used in the aging model. . .	194
9.4	Analysis of the temperature of the functional blocks for the SM with voltage assignment through PSO as compared to a baseline implementing off-chip voltage regulators.	197
9.5	Circuit and computation overhead to implement the run-time PSO. . . .	198

List of Figures

1.1	Technology trends for computing devices realized with silicon based semiconductor technology.	2
1.2	Design trends in transistor count, performance, frequency, power, and the number of cores (1970-2018).	3
1.3	The run-time, on-chip energy management techniques developed analogous to SMART Grid techniques.	6
2.1	OCVR topologies and application.	14
2.2	Schematic representation of an SCVR implementing a (a) 3:1 ladder topology, (b) depicting phase 1 operation, and (c) depicting phase 2 operation of the regulator.	17
2.3	Schematic of a multi-phase DC-DC switching buck converter.	18
2.4	Schematic of a low drop-out regulator.	19
2.5	LDO and buck converter physical area and power efficiency for moderate, high, and ultra-high switching frequencies.	22
2.6	The integrated voltage regulator module (iVRM) in IBM POWER8 processors, where (a) depicts the schematic of the iVRMs as distributed LDOs and (b) the power conversion efficiency as a function of the output voltage from the iVRM.	24
2.7	The (a) circuit schematic and (b) power conversion efficiency as a function of the load current for a fully integrated voltage regulator (FIVR) in Intel 4 th generation processors. The FIVR is a multi-phase buck converter. . .	25
3.1	NVIDIA Volta GV100 floorplan with 84 streaming multiprocessors overlaid on a map of 0.9% σ/μ variation in V_{th} for a 10 nm FinFET HKMG technology.	33

3.2	Impact of BTI on a PMOS FinFET fabricated in a 10 nm or below HKMG process through characterization of (a) the shift in V_{th} over a stress period of ten years and an operating temperature of 80 °C, and (b) the degradation in the drain to source current (I_{PMOS}) of a digital LDO with stress time and operating temperature assuming an activity factor of 50%. . . .	36
3.3	FO4 delay variation of an inverter chain with operating temperature and power supply voltage for a 7 nm PTM FinFET technology. The TEI phenomenon is depicted as an inverse dependence of the delay with temperature.	39
3.4	Reduction in the peak temperature (Joule heating) of a 15 nm FinFET technology with power supply voltage scaling.	40
3.5	Environmental and circuit effects that impact the power integrity of sub-20 nm FinFET based ICs utilizing distributed on-chip voltage regulators.	43
4.1	System model of a chip-multiprocessor considered for the disparate power management techniques. A per core local power manager (LPM) module and a global power manager (GPM) module control the power and frequency states of each core.	47
4.2	Taxonomy of power management techniques applied to chip multi-processor (CMP) systems.	50
4.3	System model of a circuit implementing adaptive voltage scaling (AVS) or voltage over scaling (VOS).	54
4.4	Simulation framework used to analyze the workload mapping algorithm.	56
4.5	Representation of the Elasticcore (a) architecture and (b) two tiered power delivery configuration with a dedicated OCVR per partition.	59
5.1	Histogram of (a) power dissipation, and (b) power variation per cycle, for the 49 SPEC CPU2000 and SPEC CPU2006 benchmarks.	69
5.2	Statistical analysis of per cycle power consumption of SPEC CPU benchmarks.	70
5.3	Proposed interconnected on-chip power delivery network with run-time voltage regulator clustering through a switching fabric.	72

5.4	Load current variation of a 16-core CMP system with 16 OCVRs each with a peak rating of 0.6 A. The stochastic model for load current consumption across the cores is based on SPEC CPU2000 and SPEC CPU2006 benchmark power trace analysis. The variation in load current is used as an input to <i>Algorithm 5.3a</i>	76
5.5	Simulated results of the implementation of <i>Algorithm 5.3a</i> . A 16-core CMP with 16 OCVRs each with a peak rating of 0.6 A is considered. The load current across cores is shown for two time stamps in (a) and (c) with the corresponding switches that are active to supply current that exceeds 0.6 A shown in, respectively, (b) and (d).	77
5.6	Startup time of the voltage regulators designed with a peak load current rating of 6A (I_{peak}) and 0.6A (I_{avg}).	80
5.7	Schematic of a multi-phase DC-DC switching buck converter.	81
5.8	Power conversion efficiency of DC-DC switching buck converters simulated with parameters listed in Table 5.7. The power conversion efficiency of the 2-phase FIVR is provided for comparison with the simulated buck converters that include enhanced light load efficiency.	87
5.9	Percentage energy saved with the proposed RPDN as compared to a baseline OCVR configuration of 6 A with enhanced light load efficiency. The energy savings is determined for the number of cores in the CMP and for the varying voltage regulator peak current rating of the DC-DC switching buck converters described in Table 5.7.	89
5.10	Structure of the reconfigurable PDN for transient analysis including (a) a schematic of a section of the PDN based on the IBM Power Grid Benchmark <i>ibmpg1t</i> with interdigitated V_{DD} and V_{SS} nets shown on different planes for the same metal layers, (b) a schematic representation of the Verilog-A model of the current limited OCVR connected to the V_{DD} nets of the modified <i>ibmpg1t</i> PDN, and (c) the outputs of two Verilog-A models of the current limited OCVR connected through a PMOS switch in an IBM 180 nm technology.	91
5.11	Transient analysis of two power distribution networks (modified <i>ibmpg1t</i> power grid benchmark) with a peak load current consumption of 25 A and 2 A served by a cluster of two OCVRs with a maximum load current rating of 15 A.	92
5.12	Comparison of the worst case voltage droop across the PDN for an unmodified <i>ibmpg1t</i> PDN benchmark and a PDN served with the interconnected current limited OCVRs combined at run time.	94

6.1	Power consumption of the Exynos big.LITTLE cores with frequency based on the model given by (6.2).	101
6.2	Power conversion efficiency of the voltage regulators serving the Exynos A15 (big) and Exynos A7 (LITTLE) cores.	102
6.3	Percentage of tasks successfully partitioned by the M-PWR heuristic and successfully scheduled by <i>Algorithm 6.3a</i> in a 16 core homogeneous CMP. Task scheduling through <i>Algorithm 6.3a</i> matches the M-PWR heuristic with the maximum output current rating of the voltage regulator set to 1 A. 107	
6.4	Contour plot of the percentage of tasks successfully scheduled by <i>Algorithm 6.3a</i> with varying taskset utilization and maximum output current of the voltage regulators in a homogeneous CMP with 16 cores.	108
6.5	Snapshot of the task assignment on a heterogeneous CMP platform with (a) big cores modeled on A15 parameters and (b) LITTLE cores modeled on A7 parameters. The maximum output currents of the voltage regulators serving each of the big cores and LITTLE cores are, respectively, 800 mA and 110 mA.	110
7.1	Projected power supply noise characterized by voltage swing normalized to a V_{DD} of 1 V for a 45 nm node.	113
7.2	Switched capacitor load circuit to emulate the switching behavior of clock edge triggered CMOS circuits. Power supply voltage is set through a PDN with characteristic impedance equal to the target impedance.	116
7.3	An impulse of load current simulated on the switched capacitor load circuit shown in Fig. 7.2. The sub-figures depict (a) the controlling bit sequence implemented as a step function, (b) the resulting impulse of load current (I_{load}), the discharging current from the decoupling capacitor (I_{CAP}), and the current through the voltage source (I_{PDN}), and (c) the on-chip voltage response to the impulse current.	118
7.4	Characterization of different grading coefficients on the (a) variation in the doping concentration on the N side of a P+/N junction, and (b) capacitance variation with the applied reverse bias voltage V_R	120
7.5	Circuit schematic of a discreet hyperabrupt varactor diode.	122
7.6	SPICE simulation of the capacitance as a function of reverse bias voltage for a hyperabrupt junction varactor based on the Skyworks SMA1211-001LF diode.	125

7.7	Voltage droop due to impulse load current generated on the rising edge of the clock. Three on-chip capacitance scenarios are simulated: ideal capacitor, SMVA-001LF hyperabrupt varactor, and custom hyperabrupt varactor ($V_J = 0.7$ V and $M=2$).	127
7.8	Characterization of a hyperabrupt varactor with variation in junction potential and grading coefficient through SPICE simulation. The bottom surface plot is from the characterization of a single varactor, whereas the top surface plot is the result of the characterization of two series connected hyperabrupt junction diodes for the topology shown in Fig. 7.9.	128
7.9	Two series connected hyperabrupt junction diodes for on-chip power supply droop mitigation.	129
7.10	Schematic of the optimized power distribution network with distributed multi-layer ceramic capacitors mounted on the PCB and package capacitors.	131
7.11	Impedance profile of the optimized power distribution network. Due to improved modeling of the voltage regulator, distributed multi-layer ceramic capacitors on the PCB, and on-package capacitors, the impedance peaks at low and mid frequency resonance are minimized.	133
7.12	Impedance profile of the power distribution network with on-chip capacitance C_{odc} and associated effective series resistance R_{odc} . The impedance profile with an ideal on-chip capacitor (blue) and a series connected SPICE model of a hyperabrupt junction diode (red) are plotted.	134
7.13	Voltage droop due to an 8 A impulse load current.	135
7.14	Voltage droop on the optimized PDN due to load activity at a resonance frequency of 50 MHz.	137
7.15	On-chip voltage and current response to resonant load activity at 100 MHz to match the impedance peak resonance of the switched capacitor load circuit. The sublets include (a) the resonant wave bit pattern controlling the switched capacitor load circuit, (b) the on-chip voltage noise for the three on-chip capacitor configurations and the current response from the load and on the PDN, and (c), (d), and (e) show, respectively, the three configurations of the on-chip capacitance; ideal, series connected Verilog-A model of hyperabrupt varactor, and SMVA1211-001LF.	138

7.16	The energy overhead of introducing additional on-chip decoupling capacitance to obtain the same amount of power supply noise suppression as series connected hyperabrupt junction diodes. Shown in the sub-figures are (a) the on-chip voltage variation as the decoupling capacitors are charged at power up, (b) the current sunk from the DC-DC regulator at power up through the optimized PDN by the on-chip decoupling capacitance, and (c) the cumulative energy consumed during power up for the overcompensated on-chip capacitance of 200 nF normalized to the cumulative energy consumed with 100 nF of series connected hyperabrupt varactor diodes. .	142
8.1	Developed voltage sense and power delivery system for 3-D ICs. A 3-D stack with device planes containing multiple voltage domains are served by a power plane consisting of multiple power modules. The ring oscillator placed in each voltage domain provides the controlling frequency to the voltage sense circuit in the power module. Two circuit techniques for the voltage sense and control mechanism are developed that provide a precise reference voltage to the on-chip voltage regulator.	146
8.2	Current starved with output switching inverter.	148
8.3	Addition based current source used as footer circuit.	149
8.4	Statistical variation of ring oscillator output frequency for 22 nm HP PTM for (a) V_{TH} , (b) effective channel length, and (c) low-field mobility. . . .	151
8.5	Block level schematic of the voltage detection and power delivery circuit for a heterogeneous 3-D IC.	153
8.6	Frequency to voltage converter circuit for detection of the power supply voltage of the device plane.	154
8.7	Transient response of the (a) ring oscillator placed in a 22 nm device plane, (b) FVC, and (c) control voltage generated from the VCVS for both a 22 nm (V_{DD_DP} of 0.8 V) and 45 nm device plane (V_{DD_DP} of 1 V).157	157
8.8	Block level view of the voltage detection and clamping circuit using a frequency comparator based topology.	158
8.9	Circuit schematic of (a) frequency detector, (b) voltage ramp generator, and (c) peak detector.	161
8.10	SPICE analysis of the power supply voltage detection and clamping circuit for a 0.8 V domain in a 22 nm technology.	163

9.1	Distributed on-chip voltage regulators for a given voltage domain. The reference voltages to the distributed voltage regulators is obtained through the on-line particle swarm optimizer. The functional units of a typical SMT processor are shown as the load circuits in the voltage domain. The timing sensors are distributed across the domain to guide the PSO. . . .	173
9.2	Assignment of a new voltage (position) to each OCVR (particle) shown through a (a) functional block diagram of the PSO algorithm and (b) a vector diagram.	174
9.3	Methodology to implement the run-time adaptive voltage assignment using the PSO as described by (a) pre-silicon design steps, and (b) post-silicon process steps.	176
9.4	A latched tap delay line in a 7 nm PTM FinFET technology used as the timing sensor where (a) depicts the circuit schematic of the sensor and (b) the output of the latches capturing the shift in the clock edges due to an increase in V_{DD}	181
9.5	Aging sensor built into each potential critical path (PCP) where (a) depicts a block diagram of the primary circuit components, (b) a circuit schematic of the sensor, and (c) a timing diagram depicting the detection of a violation due to aging.	181
9.6	Response of the timing sensor at start of life with variation in (a) V_{DD} (at $T=25^{\circ}\text{C}$ and $V_{th}=0.34\text{ V}$), (b) operating temperature (at $V_{DD}=0.7\text{ V}$ and $V_{th}=0.34\text{ V}$), and (c) intra-die V_{th} with a coefficient of variation σ/μ of 0.9 %. The distribution of V_{th} for a σ/μ of 0.9 % is provided in (d). . .	184
9.7	An illustration of determining a new voltage each time the calibration procedure is invoked due to a flag raised by an aging sensor. The assigned new voltage results in the same delay per bin in the timing sensor as observed with the previous calibration.	185
9.8	Simulation infrastructure to validate the PSO based voltage assignment algorithm where (a) depicts the floorplan of a SM used as an input to Voltspot and Hotspot and (b) the data flow diagram depicting the data parsed between the various tools used to characterize the PSO algorithm.	186
9.9	Model of the circuit used for SPICE simulation of a voltage domain that includes a circuit implementation of the PSO algorithm for voltage assignment of distributed OCVRs.	188

9.10	Results from the PSO based power supply voltage assignment algorithm for a voltage domain with 16 OCVRs. The voltage domain consists of gates from a 7 nm PTM process. The nominal voltage ν_{nom} for the transistors is 0.7 V. Included are the characterization of (a) the best voltage assignment per OCVR, (b) the global best voltage with time, and (c) the reduction in power consumption when applying the adaptive global voltage to the given voltage domain as compared to an assignment of ν_{nom}	189
9.11	Characterization of the effect of the personal (c_1) and social acceleration (c_2) coefficients on the decision of the PSO (for varying circuit activity of the functional blocks). The evolving reference voltage assignment to a sub portion of a SM with three OCVRs is shown. The change in P_{best} applied to the OCVRs supplying the FP/INT cores, register file, and tensor cores is shown for decreasing levels of circuit activity for (a) $c_1 = c_2 = 2$, and (b) $c_1 = 0, c_2 = 2$	192
9.12	The reduction in aging induced degradation of V_{th} through adaptive voltage assignment to 12 OCVRs integrated in a SM modeled on the NVIDIA Volta GV100 in a 7 nm FinFET process using (a) high- V_{th} and (b) low- V_{th} transistors. The end of life (EOL) in all cases is ten years.	196

Abstract

SMART Grid On Chip: Infusing intelligence to on-chip energy management

Divya Pathak

Advised by Prof. Ioannis Savidis

Scaling petaflop supercomputers to exascale requires a 500x increase in FLOPS, but at the cost of only a 3x increase in the total power consumption. In addition, computing systems implemented with nanometer scale multi-gate field effect transistors are increasingly integrating heterogeneous cores, GPUs, and accelerators. The power delivery and energy management of such complex chip multi-processors (CMP) is, therefore, a challenging research task spanning across the system, architecture, circuits, and device stack.

A set of on-chip energy management techniques that are similar to supply side and demand side management in a SMART grid is developed. The techniques span the circuit and system layers. Intelligence is introduced in the operation of the on-chip power distribution network (PDN) to sense variations in the computational activity across cores and reconfigure the PDN to optimize the energy efficiency. A circuit level technique that improves the energy efficiency through the implementation of under-provisioned on-chip voltage regulators (OCVRs) interconnected through a switch network is developed. An operating system level task scheduling heuristic distributes the workloads on the cores such that the required reconfiguration of the PDN is minimized. SPICE simulations indicate up to a 44% reduction in the energy consumption of the CMP.

An evolving on-chip power delivery methodology where reference voltages of the OCVRs are controlled through a particle swarm optimizer (PSO) is developed. The PSO negates the effects of transistor aging and process, temperature, and power supply noise induced variation in the load circuit, OCVRs, and on-chip timing sensors. The simulation results indicate an average reduction of 35% and 5% in, respectively, the power consumption and operating temperature of the voltage domains. In addition, the end of life of the voltage domain is prolonged due to a mean reduction in the aging induced V_{th} shift of 40%.

Novel circuit techniques to detect and set the power supply voltages and suppress power supply noise are developed. The run-time circuit techniques for power supply voltage detection and clamping are demonstrated for a heterogeneous 3-D integrated circuit through SPICE simulation of a device plane in a 22 nm technology and a power plane in a 45 nm technology. The power supply voltages of less than 1 V are successfully set and provided as a reference to an OCVR within 500 ns of initiating an active state and with variation of less than 1% in the reference voltage. Noise on the power supply is suppressed through the use of hyperabrupt p-n junction varactors as on-chip decoupling capacitors. The voltage droops and overshoots on the on-chip power distribution network are suppressed by up to 60% as compared to metal-insulator-metal (MIM) or deep trench decoupling capacitors of the same capacitance.

With approximately 42% and 15% of the data center power consumed by, respectively, the processors and the cooling system, the run-time energy management techniques developed in this thesis have significant potential to reduce the running cost of exascale data centers.

Chapter 1: Introduction

"All of the technical reports on exascale systems identify the power consumption of the computers as the single largest hardware research challenge."

– Steve Ashby, Pacific Northwest National Laboratory

With an aggressive goal of 20 MW for exaflop computing set by the U.S. Department of Energy (DOE) [1, 2], energy efficient computing across the system, architecture, circuits, and technology stack has emerged as a primary research objective. Scaling current petaflop supercomputers to exascale requires a 500x increase in computational efficiency, but with an increase in the total power consumption [1] of only 3x. The design criteria for computing devices has evolved not only based on the end application but also on technology and power constraints. The evolution of design objectives with semiconductor fabrication nodes is shown in Fig. 1.1 [3]. The primary objective is the improved energy efficiency of computing devices ranging from enterprise class systems to internet of things (IoT) components.

The first decade of the 21st century saw an end to Dennard scaling [4, 5], which was a primary driving factor of performance based scaling of computing devices built on metal-oxide-semiconductor field-effect transistor (MOSFET) technology. The increasing leakage power of MOSFET devices restricted the ability to improve performance by increasing the clock frequency of the circuit. The result has been an evolution of

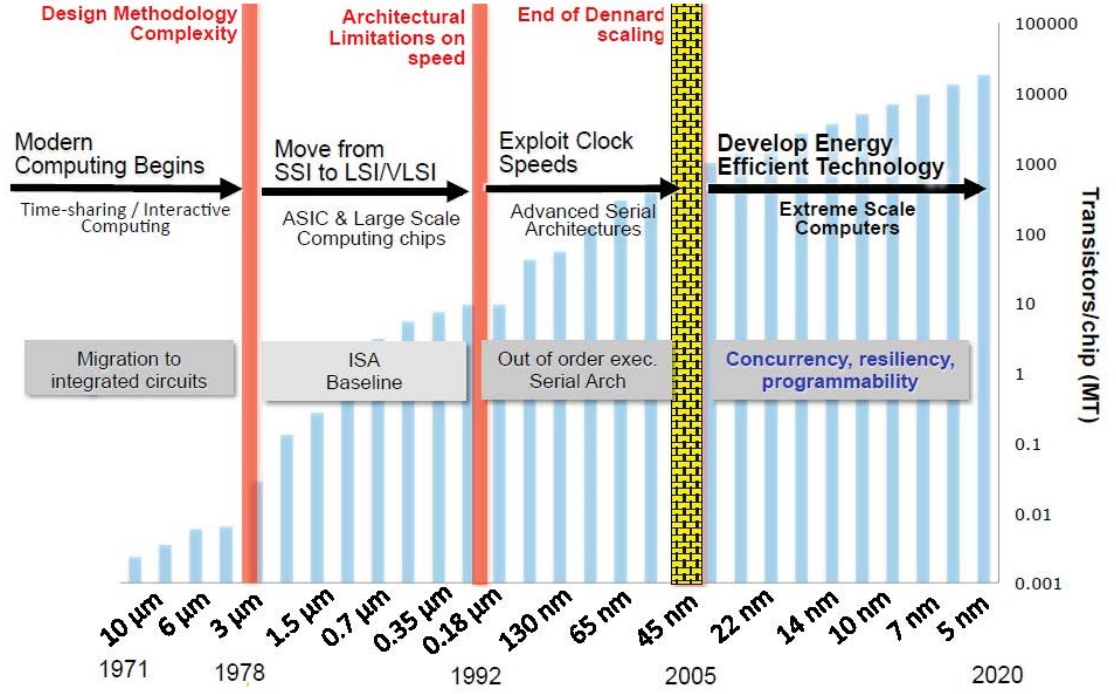


Figure 1.1: Technology trends for computing devices realized with silicon based semiconductor technology [3].

the circuit, architecture, and associated system and application software towards processor parallelism. The second decade of the 21st century is experiencing a weakening of Moore's law [6] as fabrication processes are rapidly approaching the fundamental physical limits of the transistor dimensions and there is an excessive power density when integrating a few billion transistors. The exponential rise in transistor counts attributed to Moore's law and the flattening of the curves for clock frequency, performance per clock cycle, and power consumption due to the end of Dennard scaling are shown in Fig. 1.2. As shown, the power consumption of silicon transistors has become the prime concern when designing an integrated circuit.

Systems for applications ranging from high-performance computing [10, 11], data

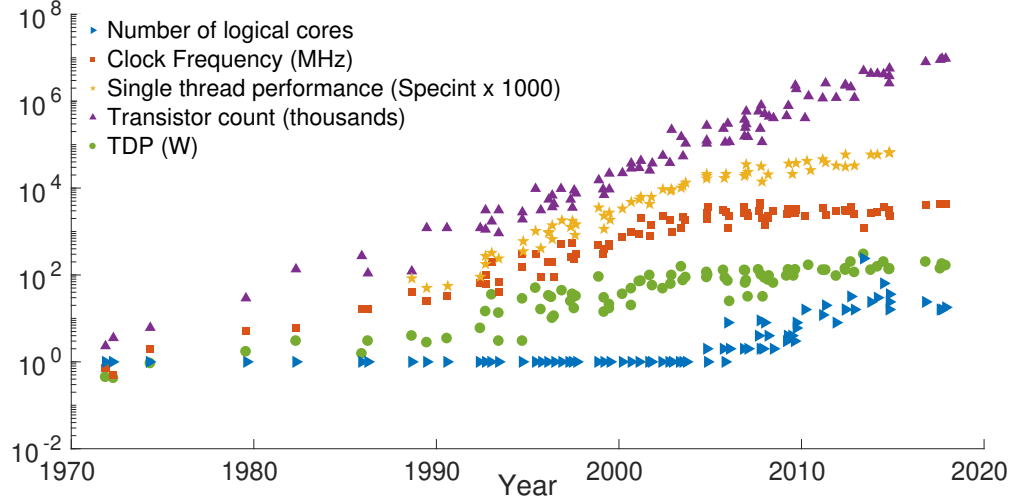


Figure 1.2: Design trends in transistor count, performance, frequency, power, and the number of cores (1970-2018). Original data up to the year 2010 is attributed to [2] and from 2010-2015 to [7]. New data is added from 2015 to 2018 [8,9].

centers [12], and neuromorphic computing [13] to low cost IoT sensor nodes are increasingly integrating heterogeneous cores and accelerators to achieve higher performance while limiting power consumption. The methods implemented by these systems to manage the power consumption and improve the energy efficiency are described in Chapter 4 along with the need for cross-layer power management, which describes techniques that span across the device, circuits, architecture, and system layers.

The analogy with SMART grid technologies for the energy management techniques developed in this dissertation is described in Section 1.1. The outline of the dissertation, research statement, and objectives are presented in Section 1.2.

1.1 SMART GRID based on-chip power management

The electric grid is amongst the largest man made systems. For instance, the U.S. electric grid consists of over 9,000 electric generation units, over 600,000 miles of transmission lines, and a generation capacity in excess of one trillion Watts [14]. The power distribution on the electric grid suffers from large economic loss due to the unpredictable consumption of electricity by the end user as well as environmental and other unforeseen disruptions. The integration of renewable sources of energy has introduced stochastic behavior in power generation. The electric grid has been modernized drastically in the last two decades due to advancements in electronic communication technologies. The modernized grid with two way information flow between the power generators and end users is described as the SMART grid [14]. Smart sensors distributed across the grid assess grid stability and monitor the pattern of power consumption by the end-users. Automated feeder switches re-route power in case of grid faults, and battery systems store sufficient energy to maintain a micro-grid that includes both renewable and non-renewable sources of generated power [14].

The power management techniques implemented in a SMART grid are classified as either supply side or demand side management. A mismatch between power demand and power supply is the primary cause of power outages and revenue loss due to over production. The unit commitment problem states that the generation of power is optimized to sustain the stochastic power demand of the system while maximizing the profit to the utility company, end user, and other economic players in the SMART grid. Switch reconfiguration to re-route electricity either to isolate a faulty

line or provide additional power to a geographic area are considered supply side management techniques. Reactive power optimization, commonly known as Voltage/VAR (voltage-ampere reactive) control, is crucial to maintain the system reliability and voltage levels throughout the SMART grid. The run-time adjustment of generator voltages, transformer taps, and shunt capacitors controls the reactive power distribution and the voltage profiles on the grid. Voltage/VAR optimization is a supply side management technique. Offsetting power consumption during the anticipated peak hours to reduce the peak load on the grid is a demand side management technique. Electric appliances in residential and industrial units are scheduled to run at off-peak hours to not only reduce the peak load but also to lower cost as electricity during off-peak hours is cheaper due to less demand.

In this dissertation a set of on-chip energy management techniques are developed that are similar to supply side and demand side management in a SMART grid. The classification of the developed techniques, which are enclosed by a dashed box, is shown in Fig. 1.3.

1.1.1 Supply side management

The primary contributions on supply side management developed as part of this dissertation are described in this subsection. A reconfigurable power delivery network with distributed on-chip voltage regulators is described in Chapter 5. Each of the on-chip voltage regulators is provisioned for a maximum output current rating meeting the average demand of the load circuit.

Hyperabrupt PN junction diodes are proposed for on-chip decoupling capacitance

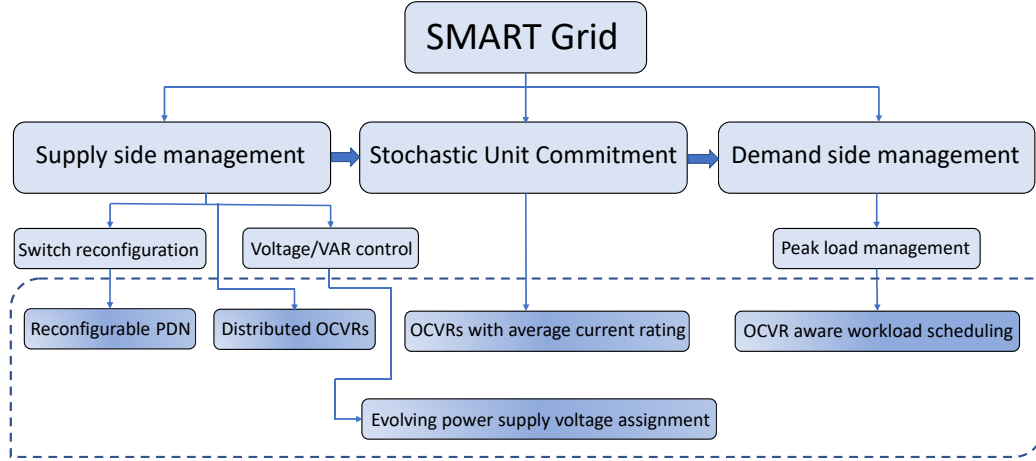


Figure 1.3: The run-time, on-chip energy management techniques developed analogous to SMART Grid techniques.

in Chapter 7. The voltage controlled capacitance offered by hyperabrupt diodes improves the suppression of clock induced power supply noise by 40% on average as compared to other state of the art on-chip noise mitigation schemes.

Circuit techniques to auto-detect and auto-set the power supply voltage of a given domain are described in Chapter 8. Although applicable to 2-D integrated circuits (ICs), the techniques are of particular use for 3-D heterogeneous ICs, where disparate dies are fabricated in different foundries and the technology specific information may not be disclosed to the die packaging facility.

An evolvable on-chip power delivery system that allows for programmability of the power supply voltage at each on-chip voltage regulator (OCVR) is described in Chapter 9. The fine grained tuning of the output voltages of the OCVRs compensates for the negative effects of aging, process variation, and power supply noise on the load circuit and OCVR.

1.1.2 Demand side management

The primary contribution on demand side management includes a workload scheduler to optimally assign tasks to the cores in a CMP. The scheduler is aware of the power delivery network and the peak current ratings of the OCVRs serving each core in the CMP. The workload scheduler described in Chapter 6 works in conjunction with the reconfigurable PDN that includes under-provisioned OCVRs, which is described in Chapter 5. The task scheduling is done such that the power demand of all the cores in a CMP is less than the power delivery capacity of the PDN. The scheduler, therefore, ensures peak load conditions do not occur on the CMP. A brief description of each energy management technique is provided in Section 1.2.

1.2 Outline of the dissertation

Existing power management solutions for many-core architectures rely on architecture level performance counters to react to the changing activity of the applications, which incurs a large penalty when applying and recovering from a low power state. Power predictive techniques are still in an early stage of development and are architecture specific. Therefore, techniques were developed that are executed at run-time through the power delivery system of a many-core system. The circuit techniques are complemented by a system level workload scheduler, which enables cross-layer energy management.

1.2.1 Research objective

The primary objective of the dissertation is to introduce intelligence into on-chip power distribution through cross layer (system and circuits) dynamic power management techniques that improve the energy efficiency of many-core systems configured with heterogeneous cores.

1.2.2 Research methods

The techniques developed for the power delivery circuits and the operating system introduce intelligence to on-chip power management to optimize the energy efficiency of the circuit while accounting for process variation, transistor aging, and power supply noise. The optimization techniques are applied at the

- operating system level through work load scheduling, and
- circuit level through the re-configurable and evolvable design of the components of the power delivery network including the voltage regulators, interconnect network, and voltage controlled decoupling capacitors.

1.2.3 Research contributions

The major contributions of the dissertation include:

1. A re-configurable on-chip power delivery network with two way information flow between the processing elements and the circuits providing on-chip voltage regulation. The peak current output of the OCVRs supports the average current requirements of the processing elements [15]. In a SMART grid, load balancing

is achieved through one or all of the following three methods: switch reconfiguration, tie-line addition, and wire upgrade [16]. In the context of on-chip power delivery, an equivalent technique to wire upgrade or tie-line addition is not feasible post fabrication. A high speed switching fabric is proposed to reconfigure the connections between the OCVRs and processing elements to perform dynamic load balancing. With optimally sized OCVRs, the energy consumption of the many-core system is reduced by up to 44% as compared to a system with OCVRs designed for the worst case. The optimal sizing of the OCVRs also reduces the on-chip foot print by at least 23%.

2. A supply side load balancing algorithm is developed for dynamic power management. The algorithm is executed on the on-chip power management unit and combines the output of the OCVRs to support load currents in excess of I_{avg} . The on-line algorithm clusters the OCVRs to provide the necessary current when the demand of the cores exceeds I_{avg} . The algorithm is an evolution of the work proposed in [17] for energy efficient OCVR clustering.
3. A convex energy optimization problem is solved to ensure the reliability of the proposed reconfigurable power delivery system with under-provisioned on-chip voltage regulators. The optimization problem is constrained by the total power budget of the IC and is limited to the combined peak current rating of the OCVRs. The feasibility of the solution, determined by solving the optimization problem, is demonstrated through a real time workload scheduling heuristic for a many-core platform. The scheduler is applicable to homogeneous and hetero-

geneous core configurations. A 100% scheduling of tasksets and assignment of DVFS levels for each core of a homogeneous CMP with task utilization factors of up to 0.55 is demonstrated.

4. Power supply noise in digital circuits induced due to sudden workload variation is a growing concern in advanced technology nodes with a sub 1 V power supply voltage and a reduced voltage guard band. Devices that offer a voltage dependent capacitance such as the silicon hyperabrupt varactor are proposed to suppress noise on the on-chip power distribution network for noise sensitive digital blocks. The variation in the capacitance of the varactor with voltage is exploited to reduce the dependence of the voltage across the varactor terminals on the charge stored or released from the varactor. For the same amount of charge drawn, the voltage drop across series connected hyperabrupt junction diodes is shown to be up to 60% less than an MIM or deep trench capacitor with the same capacitance.
5. A circuit to detect and reliably set the power supply voltage of a given voltage domain is developed. Correct power supply voltage detection and clamping is demonstrated through circuit simulation of two device planes, one in a 22 nm and the other in a 45 nm fabrication process. The power module is capable of setting the power supply voltage in the range of 0.7 V to 2.5 V. The voltage generated by the power module acts as a reference voltage to an on-chip voltage regulator. The reference voltage is within 1% of the targeted power supply voltage, as indicated by simulated results.

6. A dynamic and distributed PDN is developed through an analysis of technology parameters and the timing of critical paths of the digital blocks within each power domain. The developed method offers adaptive voltage delivery to a given voltage domain. The reference voltages of the on-chip voltage regulators are controlled through a particle swarm optimization (PSO) to compensate for the effects of transistor aging and process, voltage, and temperature induced variation. The optimized output voltages of the regulators are within the technology specified voltage tolerance bands. The on-line learning of the optimum voltages with time reduces the static voltage guard-band added during the design of the PDN for the worst case process, temperature and aging induced timing variation in digital circuits. The simulation results indicate an average reduction of 35% and 5% in, respectively, the power consumption and operating temperature of the voltage domains. In addition, the end of life of the voltage domain is extended due to a mean reduction of 40% in the aging induced shift in V_{th} .

Chapter 2: Background on DC-DC voltage converters

“According to incomplete statistics, there have been more than 500 prototypes of DC/DC converters developed in the past six decades.”

– Fang Lin Luo and Hong Ye, Essential DC/DC Converters (2005), p.1.

This chapter provides an overview of the DC-DC voltage conversion circuits used for on-chip voltage regulation. In the post-Dennard regime, the reduced noise margins in high performance integrated circuits (IC) require precise voltage delivery with low susceptibility to noise. Electronic devices ranging from mobile to server class computers traditionally include voltage regulator modules placed on printed circuit boards (PCB). In the single core computing era, the implementation of the power delivery network was less complex as one voltage regulator served the on-chip digital blocks. With the advent of multi-core computing, a single voltage regulator serving all the cores became ineffective as there is a large amount of power loss due to the varying current demands of the cores. Using multiple off-chip voltage regulators to serve individual cores significantly increases the occupied PCB area if the regulators are implemented using passive components. The slew rate offered by off-chip voltage regulators is often insufficient to meet the dynamic variation of the load circuits, leading to further power loss [18]. The increasing number of voltage domains in a System on Chip (SoC), Network on Chip (NoC), and 3-D integrated circuit require

accurate characterization to properly distribute voltage regulators to assure an optimized power efficiency. A one size fits all approach by using a single high efficiency off-chip voltage regulator fails to provide an energy efficient solution.

For the circuit and system techniques and methodologies developed in this thesis, on-chip voltage regulators are proposed to improve the energy and area efficiency of the circuit by providing point-of-load power delivery in high performance ICs. There are three important parameters that determine the quality of a voltage regulator: (1) power conversion efficiency, (2) load transient response, and (3) voltage switching time. These parameters are studied in detail and trade-offs associated with each are described in this chapter.

2.1 On-chip DC-DC voltage regulator topologies

Integrating voltage regulators on-chip to locally serve each voltage domain at a per core granularity addresses the problem of increased power losses. The IR drop is reduced as a result of: (1) the shorter interconnect length between the on-chip voltage regulators (OCVRs) and the load circuit, and (2) a smaller current is brought on chip at a higher voltage. The $L\frac{di}{dt}$ noise is better localized and reduced due to the smaller inductance offered by on-chip interconnects as compared to the large inductance of the wire bond or ball grid array when off-chip voltage regulators are implemented. The OCVRs offer a faster slew rate, generally two to three orders of magnitude higher than off-chip voltage regulators [18]. The occupied PCB area is reduced, facilitating compact form factors for mobile devices.

There are always trade-offs to consider when deciding the best voltage regulator

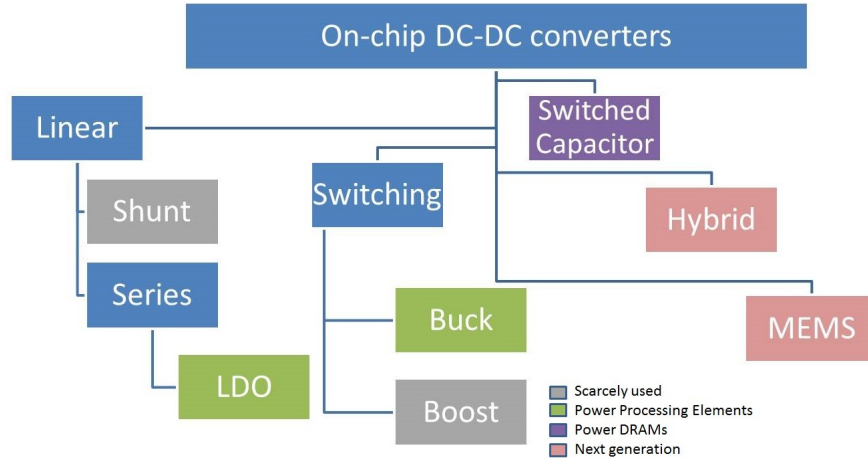


Figure 2.1: OCVR topologies and application.

topology and power distribution scheme in complex on-chip architectures. The different categories of OCVR topologies are shown in Fig. 2.1. From all topologies, switching DC-DC regulators are the most widely used due to superior voltage regulation and high efficiency. There are two important types of non-isolated switching voltage regulators: buck and boost. The boost voltage regulator is a step-up regulator that provides discontinuous output current and is intrinsically unstable due to two conjugate poles and one right half plane (RHP) zero in the duty cycle of the output transfer function. The buck converter is more widely used for on-chip voltage regulation and is described in detail in Section 2.1.2.

Linear DC-DC regulators are classified as either series or shunt topologies. Both configurations offer a step down of the input voltage. The fundamental difference between the two topologies is that a series regulator offers a path from the input voltage to the load, whereas a shunt regulator offers a path between the input voltage and ground through a variable resistor, which leads to a DC current overhead. The

lower efficiency and noise susceptibility of shunt regulators make them unsuitable for on-chip voltage regulation. Therefore, series regulators offer superior output voltage regulation and utilize minimum area as no passive components are needed. A low drop-out (LDO) regulator is a special class of series linear DC-DC regulators and is described in detail in Section 2.1.3.

MEMS based DC-DC converters [19], [20] have been proposed for on-chip voltage regulation but are far from being adopted due to process incompatibility between MEMS and CMOS and a low energy efficiency. However, once such limitations are addressed with advanced design and fabrication techniques, MEMS based DC-DC converters are potential candidates for a distributed on-chip power supply in 3-D ICs, where a dedicated plane for MEMS devices is feasible.

2.1.1 Switched capacitor voltage converters

Switched capacitor voltage regulators (SCVR) are capable of stepping up or stepping down the input voltage to any given ratio within an upper limit. The design is based on the integration of a few capacitors and switches. A single-stage SCVR is configured into several topologies based on the configuration of switches and capacitors to implement the target output voltage ratios. For instance, an SCVR implemented in a ladder-type SC converter topology is shown in Fig. 2.2. The odd-numbered switches are active in phase 1 and the even-numbered switches are active in phase 2. The DC common-mode voltage of capacitors C4 and C5 changes with respect to ground and, therefore, these capacitors are referred to as flying capacitors. The DC common-mode voltages of capacitors C1, C2, and C3 are fixed with respect to ground

and, therefore, these capacitors are termed as bypass capacitors. The conversion ratio of the ladder SCVR is 3:1 under no-load conditions.

The power losses in the resistive switches due to conduction and frequent switching limit the use of SCVRs in on-chip voltage conversion. The output voltage regulation that is achieved using the SCVR topology is inferior to other VR topologies due to: (1) the strong output voltage dependence on the current demand of the load circuit, and (2) the feedback for output voltage regulation is difficult to implement. The output voltage regulation is improved by increasing the switching frequency of the MOS switches, which requires a wider transistor width. The increased width results in greater dynamic power consumption and, therefore, a lower power conversion efficiency (PCE). Due to poor voltage regulation and PCE, switched capacitor voltage converters are not used as primary on-chip voltage regulators but are used instead to power circuits that do not require stringent voltage regulation (such as DRAMs).

2.1.2 Switching DC-DC voltage converters

Switching regulators are mixed-signal circuits that feed back an analog error signal and digitally gate the error signal to provide variable current at the output node [22]. Switching regulators offer wider output voltage levels as compared to linear voltage regulators. The general implementation of a switching voltage converter consists of a switch network and a passive low pass filter. The inductor in the low pass filter acts as a low-loss energy transfer device that enhances the PCE. For correct operation of the regulator, the switching frequency of the pulse width modulator is set to a higher value than the corner frequency of the filter. As given by (2.1), a large inductor is

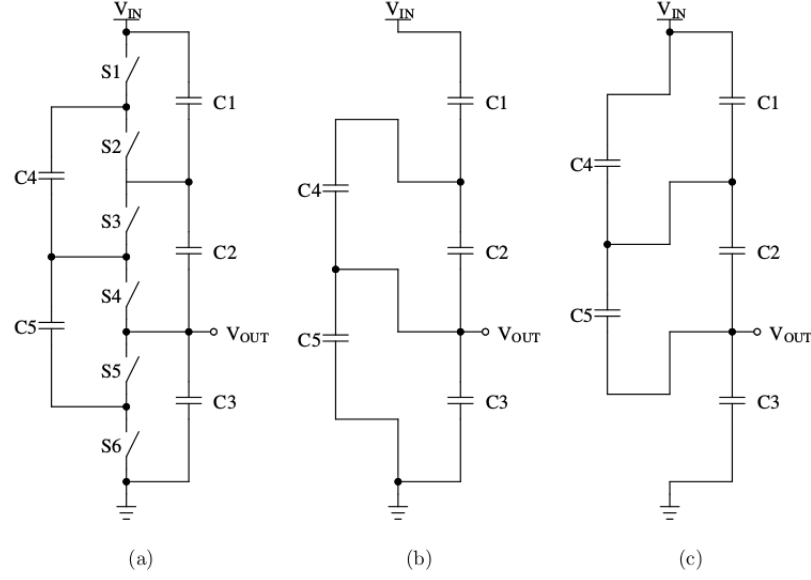


Figure 2.2: Schematic representation of an SCVR implementing a (a) 3:1 ladder topology, (b) depicting phase 1 operation, and (c) depicting phase 2 operation of the regulator [21].

implemented to ensure a low corner frequency as compared to the switching frequency of the modulator. Multiple phases of the low pass filter and the switching network are implemented to offer a wide range of output current. A schematic of a multi-phase buck converter is shown in Fig. 2.3.

The constraints in on-chip area limits the size of the inductor (usually spiral implementation), requiring the on-chip switching voltage converters to operate at high switching frequencies. However, the power efficiency of a buck converter decreases with higher switching frequency. The minimum area of a buck converter is, therefore, determined for optimum values of load current and load current ripple [23]. For a lower load current ripple, the area of a buck converter is dominated by the size of the inductor, where as for a higher load current ripple, the area is dominated by the

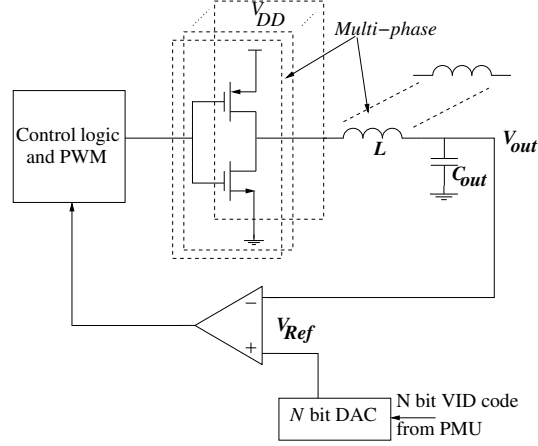


Figure 2.3: Schematic of a multi-phase DC-DC switching buck converter.

capacitor.

$$f_{cut_off} = \frac{1}{2\pi\sqrt{LC}} \ll f_{switching} \quad (2.1)$$

Most current research in on-chip buck converters is focused on achieving the best trade-off between area, power efficiency, and output power range at the chosen switching frequency [24, 25].

2.1.3 Low drop out voltage regulators

The basic structure of an LDO is shown in Fig. 2.4. The primary components include an error amplifier, a feedback voltage divider, and a series pass element implemented using a PMOS transistor as shown in Fig. 2.4. The output current is controlled by the PMOS transistor, which in turn is controlled by the error amplifier. The amplifier compares the reference voltage with the feedback voltage and amplifies the difference. If the feedback voltage is lower than the reference voltage, the gate of the PMOS device is pulled lower, allowing more current to pass and increasing the

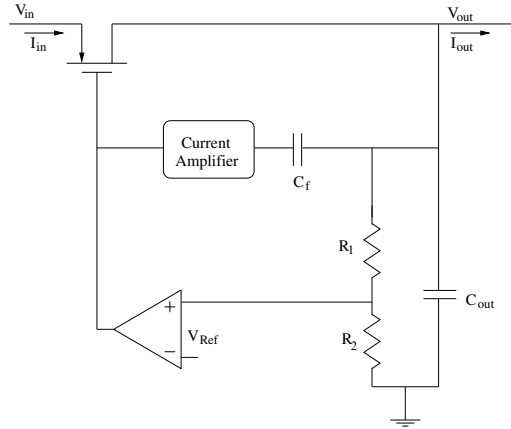


Figure 2.4: Schematic of a low drop-out regulator.

output voltage. If, however, the feedback voltage is higher than the reference voltage, the gate of the PMOS device is pulled higher, allowing less current to pass and decreasing the output voltage. As the operating current of the load or input voltage changes, the error amplifier modulates the pass element to maintain a constant output voltage. Under steady state operating conditions, an LDO behaves like a simple resistive voltage divider.

The LDO is a closed-loop system with two main poles; the internal pole of the error amplifier/pass transistor and the external pole of the equivalent series resistance (ESR) produced by the output capacitor. The power delivered to the load by the LDO is determined by (2.2), where I_{load} is the current consumed by the load circuit. The drop-out voltage V_{drop} refers to the minimum voltage drop required across the regulator to maintain output voltage regulation. The total power drawn from the power grid is given by (2.3) where $I_{quiescent}$ is the quiescent current of the LDO circuit, which characterizes the leakage current of the LDO when no output capacitance (load or decoupling) is present. The lower the drop-out voltage and the quiescent current,

the higher the achieved PCE (P_{out}/P_{in}). For on-chip voltage regulation, V_{drop} is usually small, and therefore, $I_{quiescent}$ becomes an important parameter to determine the power efficiency. A lower $I_{quiescent}$, however, results in the reduced transient response of the LDO. A trade-off is, therefore, required when designing an LDO between a fast transient response and a high power efficiency. The almost constant power supply voltage V_{IN} with technology scaling provides an upper (90%) and lower (70%) bound to the power efficiency achieved by an LDO [23].

$$P_{out} = V_{OUT}I_{load} = (V_{IN} - V_{drop})I_{load} \quad (2.2)$$

$$P_{in} = V_{IN}(I_{load} + I_{quiescent}) \quad (2.3)$$

The pass element requires a larger width to support the high load current requirement. Therefore, the total footprint of an LDO is determined by the width of the pass element. The area of an LDO in a 28 nm technology ranges from 60x60 μm^2 for a load current of 0.5 A to 150x150 μm^2 for a load current of 3.5 A [23]. The low output impedance of an LDO allows for faster load regulation and reduced voltage noise at the output node. The small foot print, low cost, low noise, and fast response to load current transients offered by an LDO makes it a suitable candidate for on-chip voltage regulation.

Table 2.1: Comparison between few on-chip voltage regulator designs [26].

Type	Buck Converter [27]	LDO [28]	Switching Capacitor [29]	Hybrid [26]
Year	2003	2007	2010	2010
Technology (nm)	80	350	32	110
Response time (ns)	87	270	NA	72-192
On-chip area (mm^2)	12.6	0.264	0.374	0.015
V_{out} (V)	0.9	1.8-3.15	0.66-1.33	0.9
ΔV_{out} (mV)	100	54	NA	44
$I_{quiescent}$ (mA)	NA	0.02	NA	0.38
I_{max} (mA)	9500	200	205	140

2.1.4 Hybrid voltage regulators

The analysis of the buck converter indicates that the footprint of the converter is dominated by the passive low-pass filter. Kose *et al.* [26] demonstrated an active filter based buck converter that does not require a passive output capacitor. The area of the converter in a 110 nm technology is 0.015 mm^2 . The response time achieved for similar output voltage requirements is faster than that of a buck converter implementing a passive filter. The characteristics of a few implementations of a dc-dc converter are compared with the hybrid voltage regulator designed in [26] and are listed in Table 2.1.

2.2 Comparison between different OCVR topologies

The increasing power density of cores requires efficient OCVRs with small area, high power conversion efficiency (PCE), and fast slew rates. Shown in Fig. 2.5 is the variation in the occupied area and PCE of the LDO and buck converter with increasing demand in the load current. For a given load current, an LDO has a smaller footprint than a buck converter. At high load currents, the footprint of an LDO becomes comparable to that of a buck converter with a high switching frequency (1 GHz). The buck converter offers lower PCE than an LDO when operating at high

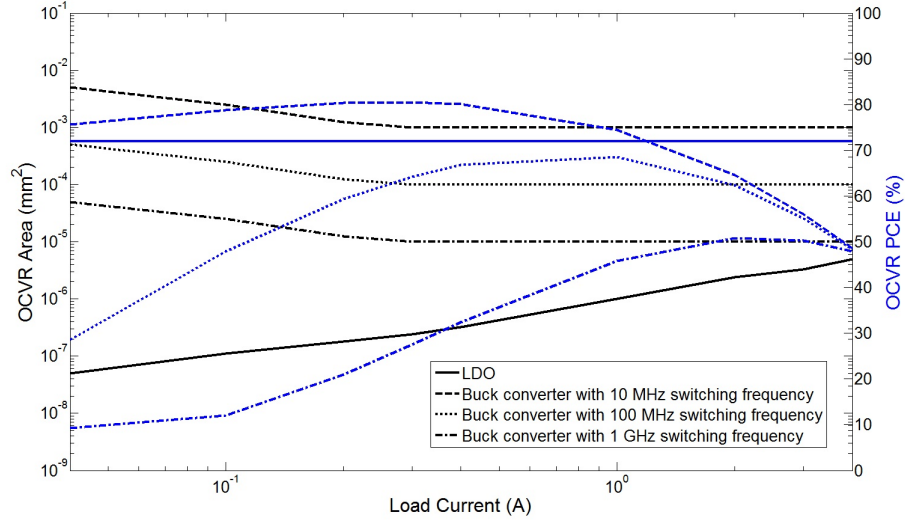


Figure 2.5: LDO and buck converter physical area, and power efficiency for moderate, high, and ultra-high switching frequencies [23].

switching frequencies. At high load currents, the PCE of an LDO is comparable to a buck converter requiring a low switching frequency. Therefore, an LDO topology offers a better PCE and occupies a smaller area but is limited by a low V_{out}/V_{in} ratio. The iVRM in the IBM Power8 processor [30] is a collection of micro LDOs that offer a peak PCE of 90.5% while supplying an output current of 11.9 A at an output voltage of 1.03 V (refer Fig. 2.6). The PCE drops almost linearly to 63% for an output voltage of 0.73 V.

Buck converters with active filters or area efficient passive filters offer a wider output voltage range with a comparable PCE as LDOs. The fully integrated voltage regulator (FIVR) in the Intel 4th generation Haswell processor utilizes inductors embedded in the routing layers of organic flip chip packaging [31]. An inductance of less than 20 nH is realized by the turns on the bottom layers of the package connected

to the on-chip power MOSFETs using plated through holes (PTH). The schematic of the FIVR is shown in Fig. 2.7a. With a switching frequency of 140 MHz and the use of a multi-phase buck converter topology, the FIVR offers a near constant PCE of 90% across a wide range of load currents, as shown in Fig. 2.7b [31]. The maximum current density offered by the FIVR is 31 A/mm², making the FIVR superior to existing on-chip LDO regulators for integration with large cores.

2.3 Challenges in implementing on-chip voltage regulators

Despite the numerous advantages of on-chip voltage regulation including reduced power supply noise, energy savings, lower IO pin count, and point of load regulation; further research in distributed on-chip voltage regulation is needed to obtain an optimal energy efficiency in high performance circuits. In this section, a few key requirements for on-chip power delivery are analyzed that offer challenging research opportunities.

2.3.1 Improvements in on-chip voltage regulator design

The voltage regulator topologies used for on-chip power-delivery are derived from the respective off-chip counterparts, and therefore, present design challenges that are not applicable to VRs placed on the PCB. The predominant limitation is the use of passive devices that consume a significant on-chip area. New integrated inductor technologies using novel magnetic materials like CoTaZr are in research for integration with silicon based technologies [32, 33]. An additional research area includes developing merged switched-capacitor multi-phase buck (MSCB) converters [34, 35].

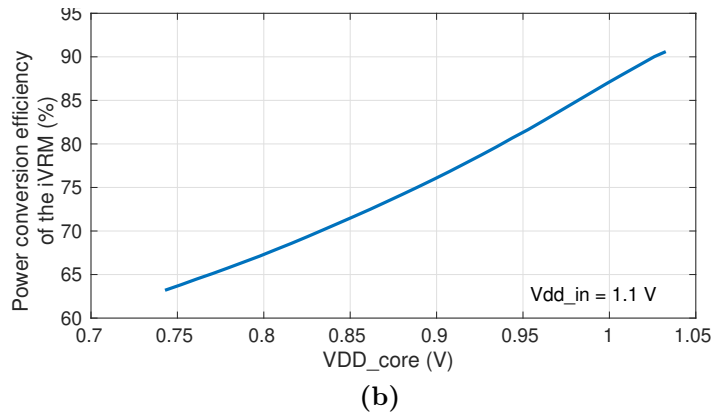
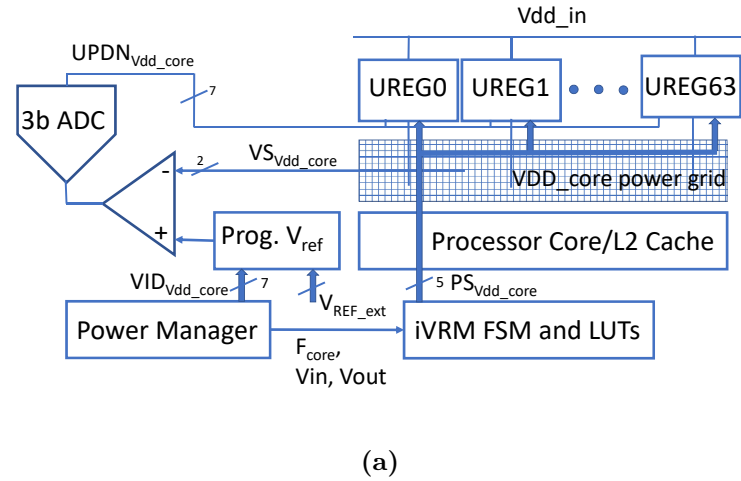


Figure 2.6: The integrated voltage regulator module (iVRM) in IBM POWER8 processors [30], where (a) depicts the schematic of the iVRMs as distributed LDOs and (b) the power conversion efficiency as a function of the output voltage from the iVRM.

Implementing a switched capacitor DC-DC converter as the first stage of voltage conversion followed by a multi-phase buck converter offers several advantages including a wider step down ratio, a reduction in area due to smaller inductors used in the buck conversion stage, and a fast dynamic response. A high power conversion efficiency is feasible across the full range of load current [25, 36].

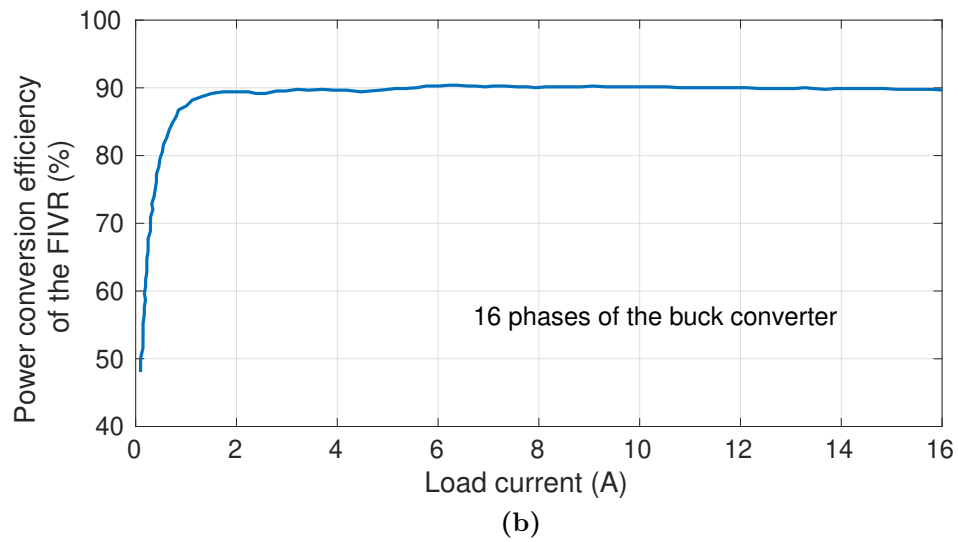
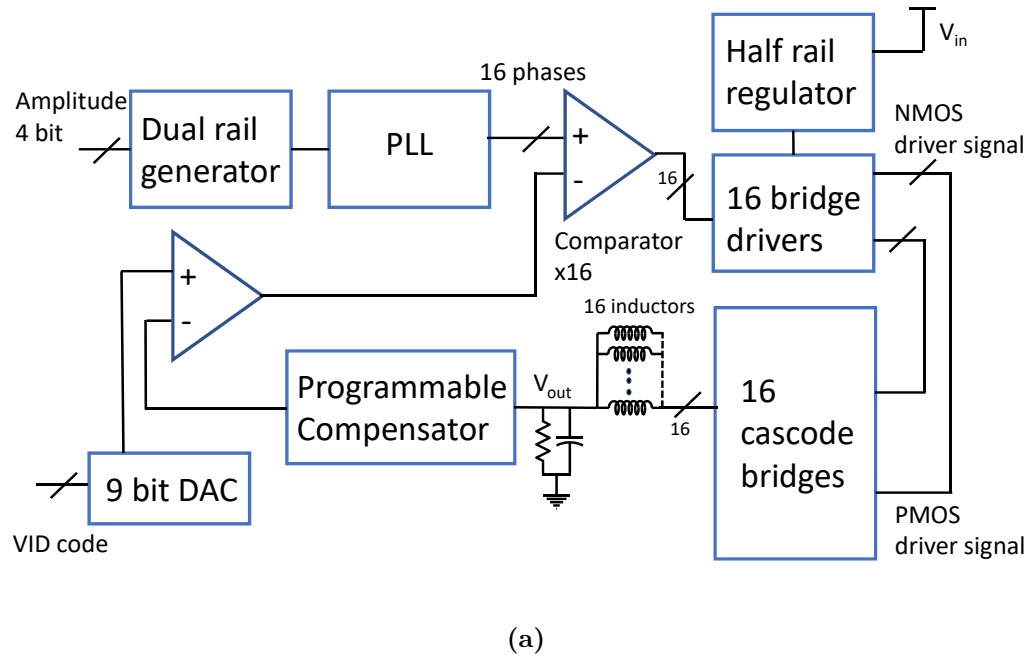


Figure 2.7: The (a) circuit schematic and (b) power conversion efficiency as a function of the load current for a fully integrated voltage regulator (FIVR) in Intel 4th generation processors. The FIVR is a multi-phase buck converter [31].

2.3.2 Distributed on-chip voltage regulation

Distributed on-chip voltage regulation is a relatively nascent research field [37], and the long term effects of on-chip voltage regulation are not fully characterized. The impact of transistor aging on the on-chip LDOs is, therefore, analyzed in Chapter 3. The IBM POWER8 and POWER9 are the first commercial processors to support distributed on-chip voltage regulation using LDO regulators [30, 38]. A common reference voltage controls the LDOs powering a given voltage domain. On-chip voltage and timing sensors guide the power management unit to generate the appropriate reference voltage for the distributed LDOs. The communication latency from the on-chip sensors increases with the on-chip area [38], which implies an increased latency when applying an updated voltage to a given power domain. A decentralized power management methodology that utilizes a *distributed intelligence* to locally optimize the operating voltage rather than a dedicated power management circuit, therefore, provides a natural evolution in on-chip voltage regulation. In Chapter 9, a method to implement distributed intelligence through swarm optimization applied to on-chip voltage regulators is proposed.

2.3.3 Adaptive voltage scaling

Dynamic voltage and frequency scaling (DVFS) is a popular technique employed in multi-core systems and SoCs to provide significant energy savings. In sub-20 nm technology nodes, the power supply voltage is scaled below 1 V offering several challenges in power integrity as highlighted in Chapter 3. Multiple voltage and frequency levels for super-threshold operation are, therefore, no longer feasible. However, adap-

tive voltage scaling (AVS) at a given operating frequency is feasible, and therefore, offers power savings at a desired performance level.

Due to the poor slew rate of off-chip voltage regulators, the performance of a digital system is reduced drastically when the voltage is scaled. On-chip regulators facilitate DVFS and AVS due to a faster transient response. A detailed study done by Kim *et al.* [18] indicates that per core DVFS using on-chip voltage regulators offers the maximum performance and the greatest energy efficiency. The energy savings is reduced as the number of voltage domains and, therefore, number of on-chip voltage regulators are reduced. The overhead in area due to the voltage regulator and associated DVFS control circuitry increases as the granularity of the voltage and frequency levels increases. There is, therefore, a trade-off between area efficiency and energy efficiency. High efficiency point of load voltage regulation with a reduced voltage switching time and a small foot-print are design constraints for OCVRs that require further research.

2.4 Summary

The increasing power density of ICs requires efficient OCVRs with a small area and a fast slew rate. The LDO topology offers the best power efficiency with the smallest occupied area, but is limited by a low V_{out}/V_{in} ratio. The buck converter implementing an active filter offers a wider output voltage range with comparable power efficiency to an LDO. An optimum topology that accounts for the number and placement of the on-chip voltage regulators requires a system level analysis where the objective is to maximize the system power efficiency while simultaneously minimizing

the overhead in power supply noise and area. Techniques such as DVFS and AVS further bolster the need for OCVRs. Improvements in CMOS circuit technologies, on-chip inductor design, and the use of novel techniques like MEMS based voltage regulators and hybrid topologies are needed to meet the competing requirements of IC power and area efficiency.

Chapter 3: Power integrity in sub-20 nm technologies

"There are two kinds of designers, those with [power]-integrity problems and those that will have them."

– on a white board at a large systems company, Signal and Power

Integrity-simplified (2010), p.1.

In this chapter, the challenges of power integrity in many core systems fabricated in advanced sub-20 nm technology nodes are described. Existing challenges including process variation, temperature related performance degradation, aging, and power supply noise are exacerbated. In addition, unique effects due to the use of FinFET technology such as self heating and temperature effect inversion (TEI) impact circuit performance. Each of the challenges and the interrelationship amongst them, which evolve with time, is described in detail for a many core system.

Physical phenomenon are described that reduce the performance of the MOS transistors in Section 3.1. Electromigration, which impacts the interconnects used for signaling and power transfer is discussed in Section 3.2.

3.1 Nanometer scale field effect transistor technology

From the invention of the field-effect transistor (FET) in 1925 by Julius Edgar Lilienfeld [39], FET technology has evolved at a rapid pace over the past eighty years. The metal oxide semiconductor FET (MOSFET) is the foundation of computing systems, which makes MOSFETs one of the primary inventions of the 20th century.

The scaling of the MOSFET to nanometer scale dimensions introduces short channel effects that degrade the carrier mobility and the drain current [24]. In addition, the voltage applied to the gate no longer fully controls the channel, leading to excessive sub-threshold leakage current, which translates to higher power dissipation [40]. To mitigate the sub-threshold leakage, two novel MOSFET structures have gained traction, the silicon on insulator (SOI) transistor [40] and three-dimensional transistors such as FinFETs [41]. SOI and FinFETs are two elegant solutions to simultaneously maximize gate-to-channel capacitance and minimize the drain-to-channel capacitance of a MOSFET.

The SOI MOSFET has a buried oxide layer, which isolates the body from the substrate. The isolation of the transistor body from the substrate reduces the parasitic capacitance and the leakage current to the substrate. Fully depleted SOI (FD-SOI) devices have very thin (5 nm to 20 nm) body structures [40]. FD-SOI technology offers superior sub-threshold characteristics with low leakage current as well as a reduced drain to source capacitance, which results in a reduction in the delay and dynamic power consumption of the transistor [40].

The FinFET structure was developed by Hu et.al in 1999 [41]. A double gate structure was proposed to enhance the control of the channel. The current three dimensional structure of the FinFET has a thin vertical fin of silicon surrounded by the gate on either two or three sides [42]. The effective width of the channel is, therefore, a function of the fin height. To provide greater drive strength, the fin height is increased. Alternatively, multiple fins implemented in parallel and connected through a common gate also offers a higher drive strength.

For nodes smaller than 22 nm, semiconductor companies have adopted either SOI or FinFET technologies due to the various design and manufacturing trade-offs. Although SOI offers less manufacturing complexity due a compatibility with bulk MOSFET, the cost of the SOI wafer is higher than that of a FinFET wafer [43]. In addition, FinFETs offer higher drive current as compared to SOI.

In the subsequent sections, the primary and second order effects that impact the power integrity of FinFETs are described. With the on-chip implementation of the voltage regulators, the existing methods to improve the power integrity of sub-20 nm nodes are insufficient. This chapter, therefore, highlights the need to consider the combined effect of all technological characteristics and challenges that degrade the power integrity of FinFETs in both the load and power delivery circuits.

3.1.1 Process variation

The stochasticity of the manufacturing process of an integrated circuit is a well researched field [44]. Each new process node developed by a foundry requires detailed analytical and experimental validation to quantify the variation in parameters due to the fabrication process. The effect of process variation is addressed by adding either (or both) timing margins and voltage margins to respectively, the maximum operating frequency and the minimum operating supply voltage. The margins also account for the predicted degradation in circuit performance due to aging over the lifetime of the IC. The use of margins limits the benefits in power and performance possible with technology scaling.

An analysis of the required power supply voltage of GPU streaming multi-processors

(SMs) for a 5% variation in the ratio of the standard deviation σ over the mean μ (σ/μ ratio) for the threshold voltage V_{th} is performed in [45]. Results of the study indicate that regions of the GPU most susceptible to process variation must operate at a 25% higher voltage than regions robust to variation. A similar analysis of the NVIDIA GV100 GPU is performed. The GV100 die size is 815 mm² with over 21 billion transistors fabricated in a TSMC 12 nm FinFET process [46]. There are 84 SMs on the die. Assuming a similar floor plan and die configuration in a 10 nm HKMG process [47], the variation in the V_{th} across the die is shown in Fig 3.1. A modest 0.9% σ/μ ratio with a spatial correlation range ϕ of 0.2 is assumed [48]. As indicated by results shown in Fig. 3.1, there is significant inter and intra SM variation in the threshold voltage. The GV100 does not support on-chip voltage regulation and, therefore, includes twelve off-chip voltage regulators. The structural complexity of the PDN for the different voltage domains is not provided [49]. Given the possibility of high variation in the operating voltage and frequency of the SMs, distributed on-chip voltage regulation with run-time power management is needed for state of the art GPUs as well as other many-core architectures.

3.1.2 MOS transistor aging

The performance of MOS transistors is effected by physical phenomenon. Specifically, bias temperature instability, time dependent dielectric breakdown, and hot carrier injection impact a circuit as it ages. In literature, these phenomenon are investigated in isolation. The impact on the on-chip DC-DC voltage regulators and the load circuits due to MOS transistor aging is, therefore, analyzed in Chapter 9.

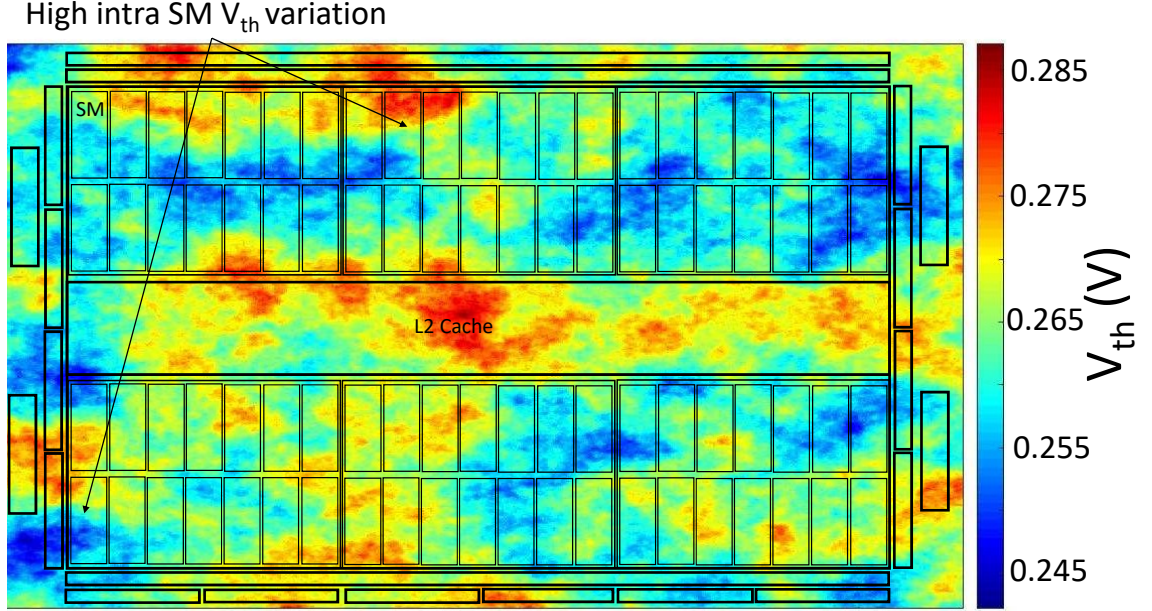


Figure 3.1: NVIDIA Volta GV100 floorplan with 84 streaming multiprocessors [49] overlaid on a map of 0.9% σ/μ variation in V_{th} for a 10 nm FinFET HKMG technology [47].

Bias temperature instability

Among the multiple reliability issues, transistor aging due to negative (positive) bias temperature instability NBTI (PBTI) is a primary failure mechanism. NBTI (PBTI) has emerged as the dominant aging effect in advanced technology nodes below 65 nm. A negative (positive) voltage applied to the gate of the PMOS (NMOS) transistor results in an increase in the threshold voltage V_{th} , which degrades the drain current, and therefore, the transconductance of the PMOS (NMOS). The phenomenon is termed as NBTI (PBTI). MOSFET circuits, therefore, exhibit a degradation in the delay with time, with more than 20% degradation due to BTI reported in literature [50–52]. The effect of BTI on a circuit depends on several factors including operating

temperature, workload activity, applied voltage, and the total active time of the circuit. The BTI effect significantly reduces the lifetime of a CMOS transistor. With time, the critical path delay exceeds the timing constraint, which leads to timing failure.

Two mechanisms contribute to the gradual increase in the threshold voltage of the device: 1) Weak Si-H bonds at the Si- O_2 interface break due to the high vertical electric field. The break of the bond releases hydrogen atoms that diffuse into the gate oxide leaving an interface trap [53]. The generated traps capture charge carriers that have tunneled into the oxide. (2) Charge carriers are captured via tunneling in pre-existing defects at the gate oxide interface with the silicon or within the gate oxide itself [54]. When the device is turned off, some of the activated defects are annealed, which leads to partial recovery of the threshold voltage. Current methods to characterize and address the BTI effect include 1) analysis and modeling of BTI in the pre-silicon phase [55], 2) input vector control and power gating techniques [56], and 3) run-time techniques such as computational sprinting at an elevated power supply voltage followed by power gating to recover from BTI [57, 58].

The NBTI induced degradation in the V_{th} of a PMOS transistor in a 10 nm HKMG process is shown in Fig. 3.2(a). An operating temperature of 80°C and a duty cycle of 50% are assumed to compute the variation in the mean and standard deviation of the V_{th} [47, 59]. Not only does the mean of the V_{th} increase due to NBTI, the variance at the start of life attributed to process variation also increases with time and operating temperature. The increase in the V_{th} results in a reduction in the maximum operating frequency of the processing elements (the cores or circuit

sub-blocks) and also a degradation in the line and load regulation offered by the distributed OCVRs. The percentage reduction in the drain to source current of the header PMOS (I_{PMOS}) in a conventional LDO is shown in Fig. 3.2(b). The reduction in the performance of the load circuit and the degradation of the maximum load current supported by the OCVRs differ due to disparities in stress time, spatial variation in V_{th} and temperature, and transistor duty cycle. Therefore, a design time model of aging induced variation is not accurate to compensate for the run-time performance loss of a system operating at a fixed power supply voltage.

Time dependent dielectric breakdown

Time dependent dielectric breakdown (TDDB) is a phenomenon that occurs due to the wear out of the gate dielectric with time. The degradation of the gate dielectric results in a conducting path through the transistor gate, which leads to a permanent short. With an increase in the operating voltage and temperature, the mean time to failure (MTTF) decreases. The mathematical model that best approximates the MTTF due to TDDB is given by (3.1) [60]. The MTTF is a function of the gate to source voltage V_{gs} applied to the transistor, the operating temperature T , the duty cycle D , and the fitting parameters a , b , X , Y , and Z [60]. The fitting parameters are constants with values $a = 78$, $b = -0.081$, $X = 0.759\text{ev}$, $Y = -66.8\text{evK}$, and $Z = -8.37\text{E-}4\text{ev/K}$ provided for a reliability aware microprocessor model. The $MTTF_{TDDB}$ from the model exhibits exponential degradation with temperature. Due to the 3-D fin structure, multi-gate FinFETs have lower heat dissipation, leading to reduced

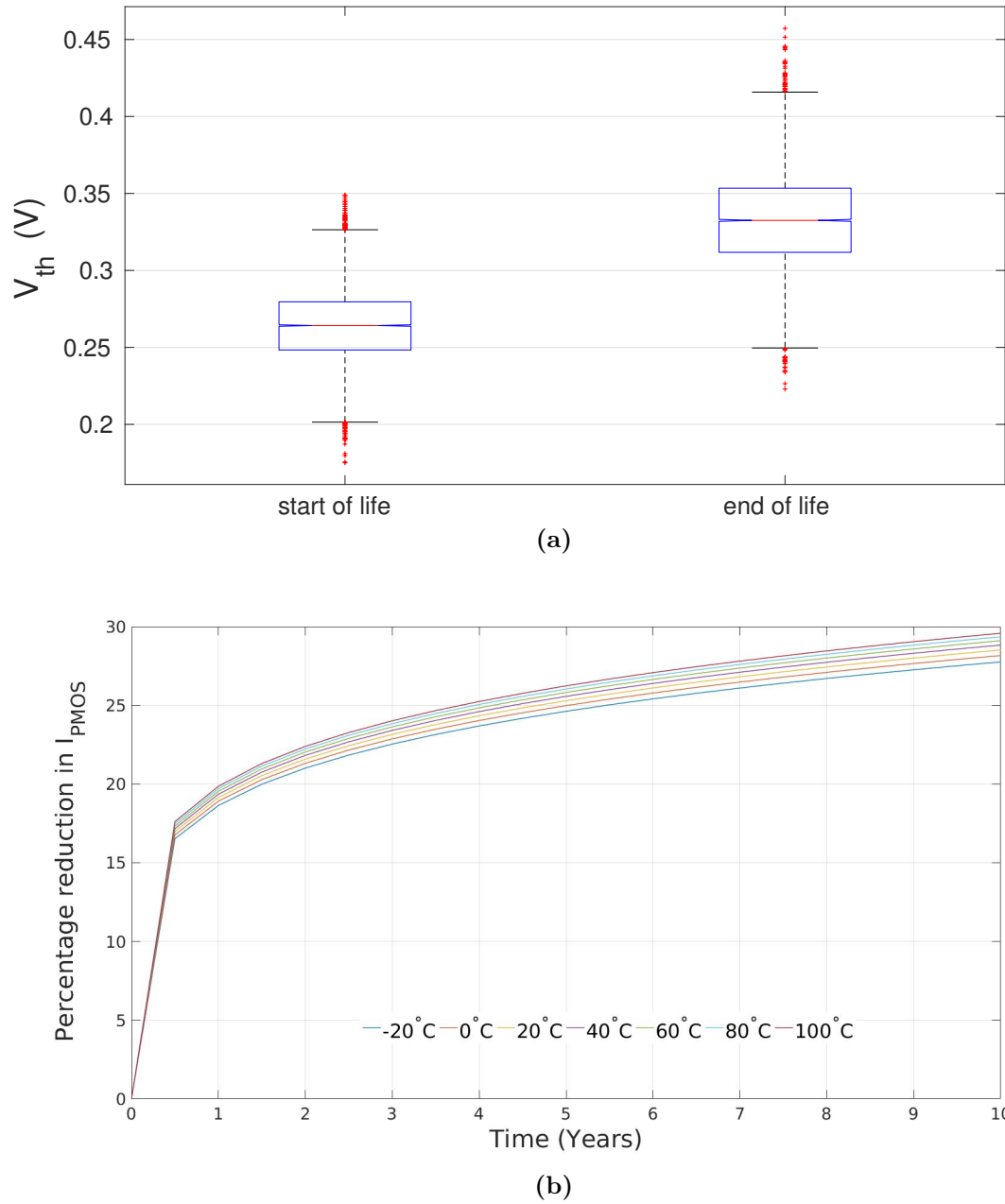


Figure 3.2: Impact of BTI on a PMOS FinFET fabricated in a 10 nm or below HKMG process [47, 59] through characterization of (a) the shift in V_{th} over a stress period of ten years and an operating temperature of 80 °C, and (b) the degradation in the drain to source current (I_{PMOS}) of a digital LDO with stress time and operating temperature assuming an activity factor of 50%.

MTTF.

$$MTTF_{TDDDB} \propto (V_{gs}^{(-a+bT)} \times e^{\frac{X+Y/T+ZT}{kT}})^{-1} \quad (3.1)$$

Hot carrier injection

Hot carrier injection (HCI) occurs when electrons are accelerated by a high electric field in the channel of a device such as near the drain of a MOSFET. The accelerated electrons, termed as hot, damage the gate oxide, which results in trapped charges. With time, the accumulation of charge due to hot electrons leads to an increase in the threshold voltage of the MOSFET. HCI is a widely studied phenomenon in bulk CMOS devices and persists in sub-20 nm FinFET devices. Studies on FinFET transistors indicate significant degradation in NMOS characteristics due to HCI as compared to PBTI [61].

3.1.3 Transient faults

Challenges in signal and power integrity include single and multiple transient faults that occur during the operational lifetime of a circuit. There are two categories of transient faults: 1) radiation effects (single event transient/upset) and 2) random telegraph noise (RTN). Research has shown that technology scaling results in increased susceptibility to single event upsets due to radiation, though the single error rate does not increase significantly [62]. RTN, however, is considered the primary transient noise source as technology scales [63]. The cause of RTN is the trapping and de-trapping of charge carriers at the interface of the silicon with the gate insulator, which is an intrinsic quantum process [64]. A single trapped charge

carrier can lead to RTN. With technology scaling, the thickness of the gate insulator of MOS/FinFETs is reduced to 1 nm [64], which enhances the direct tunneling of currents from the channel to the traps in the oxide. Therefore, quantum effects such as RTN are more pronounced in sub-20 nm nodes where the gate insulator thickness is reduced.

3.1.4 Temperature effect inversion and self heating in FinFETS

The delay of planar transistors increases with temperature due to a decrease in the drive current capability of the device. FinFETs exhibit a reverse trend in delay with temperature. In FinFET based circuits, the decrease in gate delay with temperature is described as temperature effect inversion (TEI) [65] as the drive current of the FinFET strengthens with increasing temperature. The increase in the drive current is due to the tensile stress effect from the insulator layer to the body of the vertical fin, which affects the carrier mobility [65]. The phenomenon is more pronounced in technology nodes below 22 nm. SPICE simulation of an inverter chain is performed on a 7 nm predictive technology model (PTM) of a FinFET process [66] to determine the variation in the F04 delay with temperature and power supply voltage, with results shown in Fig 3.3. The delays are normalized to a nominal voltage of 0.7 V. Similar to FinFET devices in 10 to 20 nm processes [67], the 7 nm PTM transistors also exhibit TEI, with greater variation in delay at sub/near threshold voltages.

The vertical fins are embedded in an oxide layer with a very narrow connection between the fins and silicon body. Due to the high current densities in ICs fabricated

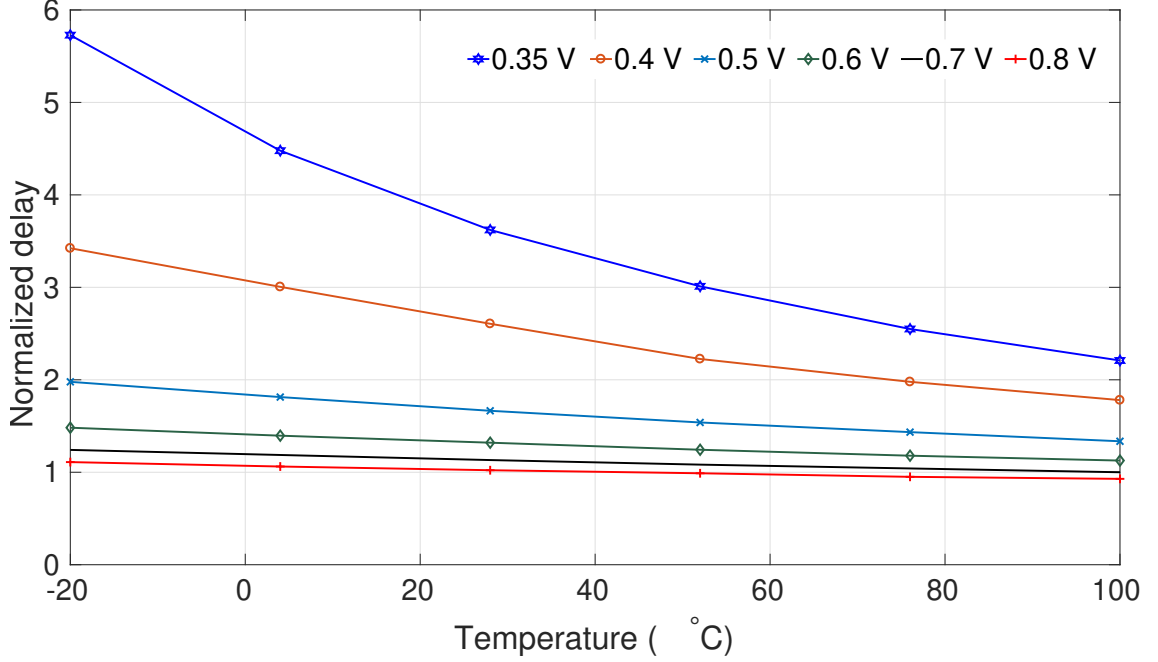


Figure 3.3: FO4 delay variation of an inverter chain with operating temperature and power supply voltage for a 7 nm PTM FinFET technology [66]. The TEI phenomenon is depicted as an inverse dependence of the delay with temperature.

using sub-20 nm FinFET based transistors, the operating temperature on the fins rises due to the poor heat dissipation to the silicon bulk. The localized temperature of central fins in a multi-fin transistor is higher due to an increase in the distance from metal contacts. The localized heating due to high current densities and poor heat dissipation of FinFETs is described as the self heating effect (SHE). Prior work has shown that degradation in reliability due to aging (BTI, HCI, and TDDB) is accelerated by SHE [68]. The models for BTI, HCI, and TDDB increase in complexity when accounting for SHE, as SHE is a localized phenomenon impacting individual FinFETs rather than circuit blocks. Therefore, it is a challenge to mitigate the effects of SHE with conventional techniques to reduce hotspots. In addition, the SHE of FinFETs

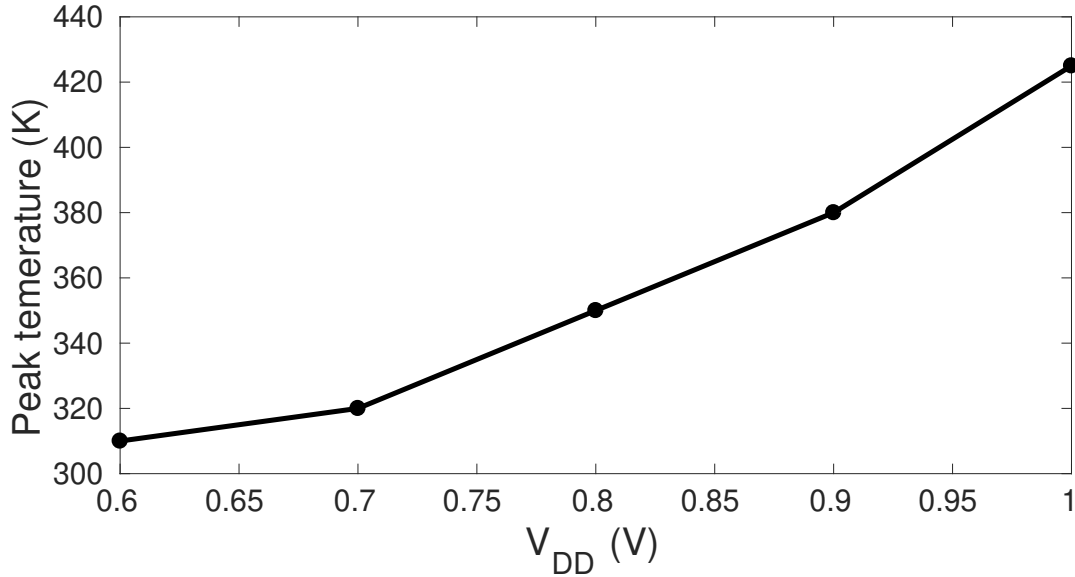


Figure 3.4: Reduction in the peak temperature (Joule heating) of a 15 nm FinFET technology with power supply voltage scaling [69].

exacerbates the effect of electromigration on the metal interconnects surrounding the device. Both power rails and signal interconnects are affected due to self heating. Prior work [69] has shown significant reduction in the lattice temperature with power supply voltage scaling. Therefore, as indicated by Fig. 3.4, power supply voltage scaling provides a significant reduction in the SHE of nanoscale FinFETs. However, existing on-chip sensor based adaptive voltage techniques or thermal management techniques do not account for the complex interaction of SHE with BTI, HCI, and TDDB as well as the inverse relationship of the FinFET delay with the operating temperature (TEI).

3.2 Electromigration

Scaled transistor technologies have also exacerbated the probability of electromigration (EM) in interconnects used for signaling and the power distribution network. The fundamental causes include increased current density, a tight-pitch between interconnected metal layers, self-heating in 3-D multi-gate transistors, and increasing interconnect resistance. In addition to the reduction in the interconnect cross-sectional area, carrier scattering from the boundaries of individual metal crystal grains is contributing to an increase in the interconnect resistivity [70]. EM is difficult to predict and prevent through chip or wafer testing as a correlation exists between various parameters including manufacturing defects, workload, and environmental conditions.

3.3 Combined effect of PVT, aging, SHE, TEI, and noise

The physical phenomenon related to trapped charge carriers at various interfaces of the MOS/FinFET lead to a degradation of the V_{th} and, therefore, a reduction in the characteristic operating frequency of the transistor with time. The scaling of the transistor results in an increased sensitivity to charge trapping and de-trapping. The failure probability of the interconnect due to EM also increases. The experimental and physical models of BTI, TDD, HCI, and EM indicate a direct relationship with the operating voltage of the transistor. This direct relationship offers an opportunity to apply adaptive voltage scaling (AVS) during the lifetime of the circuit to reduce the rate of degradation in circuit parameters due to aging. A method to apply AVS to improve the power integrity of circuits fabricated in sub-20 nm nodes is developed

and described in Chapter 9.

The disparate effects of PVT variation, aging, SHE, TEI, and noise that impact the power integrity of high performance integrated circuits fabricated in sub-20 nm FinFET nodes have thus far been addressed individually while developing power or thermal management solutions [67,71,72]. Design time modeling of process variation, EM, BTI, HCI, TDDB, RTN, TEI, SHE, and IR drop on the PDN are insufficient to predict the combined effect on the load circuits, the on-chip sensors that monitor the circuit properties, and the on-chip power distribution network including the distributed voltage regulators. The on-chip components affected by the various phenomena are shown in Fig. 3.5. The package and the global power distribution network are not shown in Fig. 3.5, but are also impacted by EM, power supply noise, and transient faults. The combined effect of aging induced variation on the OCVRs (different topologies age at different rates when subject to the same load, temperature, and environmental conditions) and the load circuit is unknown. As shown in Fig. 3.2(b), the PMOS header in the low drop out regulator experiences a reduction in the drive current I_{PMOS} with temperature, circuit activity, and time, due to an increase in the threshold voltage. The degradation in the drive current of a buck converter with age differs from an LDO due to the complex function of the MOS power switch, MOS drivers, and the pulse width modulation circuit. In addition, such model based aging analysis ignores the passive components and the interconnects in the OCVR that suffer from increased EM as a function of the duration of the applied stress.

The combined effect of aging and temperature on the interconnects and sub-20 nm FinFETs has not been modeled in literature. The parasitic resistance and

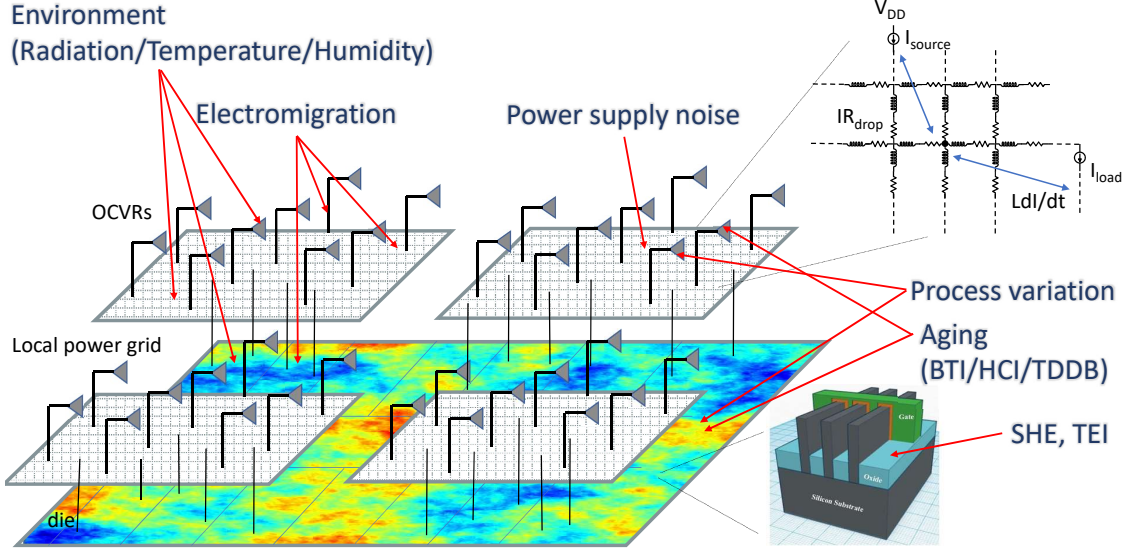


Figure 3.5: Environmental and circuit effects that impact the power integrity of sub-20 nm FinFET based ICs utilizing distributed on-chip voltage regulators.

capacitance of the interconnects increase with temperature; however, the current drive of the FinFET transistors improves due to TEI, which reduces the delay of the path. Alternatively, given disparate stress times for the load circuits and the spatial temperature variation across the die, the rate of aging in the load circuit and the OCVRs is not identical. In a circuit with distributed OCVRs, assigning the same reference voltage V_{ref} to all the OCVRs is not the best practice as process variation and an aging induced shift in V_{th} affects the load circuits and the OCVRs non-uniformly, even in the same voltage domain. Due to the complexity of implementing disparate V_{ref} circuits for a large number of distributed OCVRs, a tradeoff between the accuracy of the assigned local voltages and the circuit cost (area and power) of the implementation is required [37].

Due to the limitations of accurately modeling or predicting the impact of the dif-

ferent phenomenon that degrade the power and signal integrity of a deeply scaled system, the use of voltage guard-bands appears to provide the most effective solution for advanced technology nodes. However, given the highly scaled operating voltages and diminishing difference between super threshold and near threshold operation, adding guard-bands to the operating voltage negates the benefits of technology scaling with regard to reducing the power consumption of the circuit. Timing guard-bands that are added to the critical path(s) of the circuit are, therefore, a more conservative solution while resulting in a reduction in the performance benefits of sub-20 nm FinFET technology. By definition, a critical path in a synchronous clocked digital system is the path with the largest signal propagation delay. The maximum clock frequency is set by the critical path delay. Conventionally, critical paths are identified during the statistical static timing analysis (SSTA) of the sequential circuit. However, in advanced technology nodes, there are multiple challenges to correctly identify all the critical paths in a circuit. The critical path is dependent on the executing workload, which is a run-time variable unknown during SSTA. Process variation and aging effects, which include both spatial and temporal variation, influence the delay of the paths. Therefore, new critical paths emerge with time [73]. Due to the complex interaction of the workloads, process variation, ambient conditions, and circuit aging, critical path identification is no longer deterministic at design time. Therefore, a run-time technique is needed to identify evolving critical paths in the circuit and auto-adjust the operating voltage and clock frequency to mitigate timing failures on the critical paths.

3.4 Summary

In this chapter, the impact of process variation, transistor aging, temperature, and environmental effects are investigated for processing elements such as the GPU and on-chip voltage regulators fabricated in sub-20 nm FinFET technology. Despite the negative impact on circuit characteristics due to these second order effects when scaling, a holistic solution that ensures the power integrity of the integrated circuits fabricated with sub-20 nm FinFET technology is not addressed in literature. A run-time power management solution that *learns* and assigns the local voltages across the IC through distributed OCVRs is needed to ensure the power integrity of the circuit while minimally impacting the energy efficiency and performance. The optimum online learning algorithm for run-time voltage assignment in high performance many-core circuits is, therefore, developed in Chapter 9.

Chapter 4: On-chip power management

"If the Cloud were a country, it would have the fifth largest electricity demand in the world."

– Garry Cook, Greenpeace International

In this chapter, an overview of popular power management techniques published in literature, and those supported in commercial single and multi-core processors, is provided. The techniques are categorized as those implemented during the design of the circuit and those executed at run-time. Although executed at run-time, the techniques require the inclusion of circuits during the design of the processors. The generic system model considered for the power management techniques is shown in Fig. 4.1. The local power manager (LPM) is responsible for configuring the power modes at the core level [74]. The global power manager (GPM) is distributed across the operating system, architecture, and circuit layers [74–76]. The GPM controls the core parameters such that the total power consumption across all cores does not exceed a limit set by the predefined thermal budget. The two primary parameters tuned by run-time techniques are the operating voltage and frequency of each core. The primary performance counter employed by the LPM and/or the GPM is the instruction per second (IPC) information. The disparate power management techniques target reductions in either (or both) the dynamic and static power consumption of the cores.

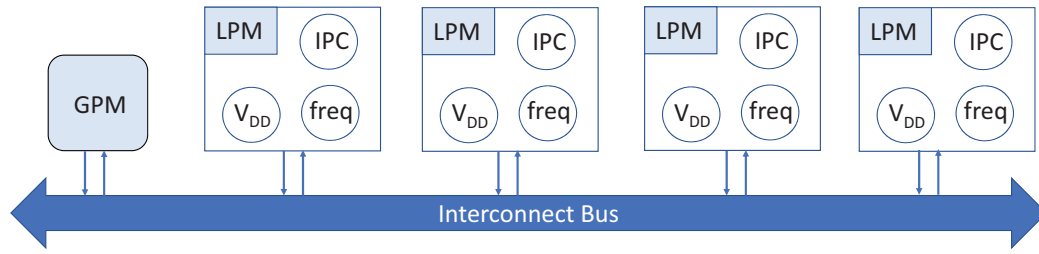


Figure 4.1: System model of a chip-multiprocessor considered for the disparate power management techniques. A per core local power manager (LPM) module and a global power manager (GPM) module control the power and frequency states of each core.

The taxonomy of power management techniques applied to chip multi-processors is shown in Fig. 4.2. The taxonomy is restricted to techniques applied to the processing cores. The power management techniques for the different levels of caches (SRAM), main memory (DRAM), and high speed interconnect buses (PCIe, NVLink etc.) are not considered. The power management techniques for the core are applied at the technology, circuit, micro-architecture, and operating system layers of the computing stack. Many techniques span across the four layers including existing run-time power management methods.

4.1 Technology level power management

In addition to technology scaling, three-dimensional transistors with multiple gates such as FinFETs and gate all around (GAA) devices have been developed for advanced technology nodes by many leading semiconductor foundries. Below 22 nm, planar transistors suffer from undesirable short-channel effects that result in the increased leakage current of the transistors as described in Chapter 3. Since the fourth

generation of microprocessors, Intel has fabricated circuits with Tri-Gate or FinFET transistors [77]. In addition to reduced leakage current and better control of the on-state channel current, multi-gate transistors are much more compact than planer transistors, which results in an increase in the integration density [78].

4.2 Circuit level power management

At the circuit level, the predominant techniques to reduce power include transistor sizing [79], low swing logic [80], low power clocking [81], and threshold voltage V_{th} optimization [82]. V_{th} optimization must be supported by the foundry, and is typically applied through the technology library. During circuit synthesis, the selection of low- V_{th} or high- V_{th} devices is dependent on the design objective of either, respectively, high performance or low power. The low- V_{th} devices switch faster but suffer from higher leakage current, and the high- V_{th} devices suffer from higher switching latency but reduce the power consumption due to leakage. The synthesis tools perform multi- V_{th} optimization based on a fixed leakage power budget.

Near-threshold (NTC) and sub-threshold (STC) circuits have gained traction in the past decade for low power, battery operated systems. For neuromorphic computing, sub-threshold circuits are preferred due to the exponential relationship exhibited between the drain current and the applied gate-to-source voltage. The exponential relationship more accurately represents the response of a neuron as a transfer function that maps the input current into the frequency of the spikes generated by the neuron. Approximate computing is another field that makes use of circuits operating in sub- V_{th} . The circuit techniques and methodologies to operate at supply voltages

near and below the threshold voltage of the transistors requires a trading-off between performance, error-resilience, and low dynamic power consumption. Although the leakage energy does increase exponentially at sub- V_{th} voltages, multigate technologies limit leakage current. The circuits are, therefore, set to a voltage that optimizes the energy consumption while maintaining a tolerable error rate and performance.

4.3 Architecture level power management

Techniques for low power architecture design predominantly include adaptive processor architectures [83], power gated and clock gated designs [84], domino logic [85] and low power control logic design [86]. An adaptive processor architecture activates a minimum set of circuit blocks to execute the given operating code. Examples include adaptive instruction issue queues [87] and adaptive caches [88]. In adaptive instruction queues, the circuit partitions that contain currently executing instructions are powered and the remaining blocks are deactivated. Similarly, in an adaptive cache, the lines, blocks and sets containing data are selectively powered. A clock gated or power gated architecture selectively removes the clock signal or the power source to a non-active functional unit in the processor. Clock gating reduces the dynamic power consumption of the circuit whereas a power gated block limits both the dynamic and leakage power consumption. However, when considering energy efficiency and performance, power gating includes added overhead due to the inrush current and the additional clock cycles needed to wake a power gated block. The inrush current also increases the peak voltage noise on the power network of the gated blocks.

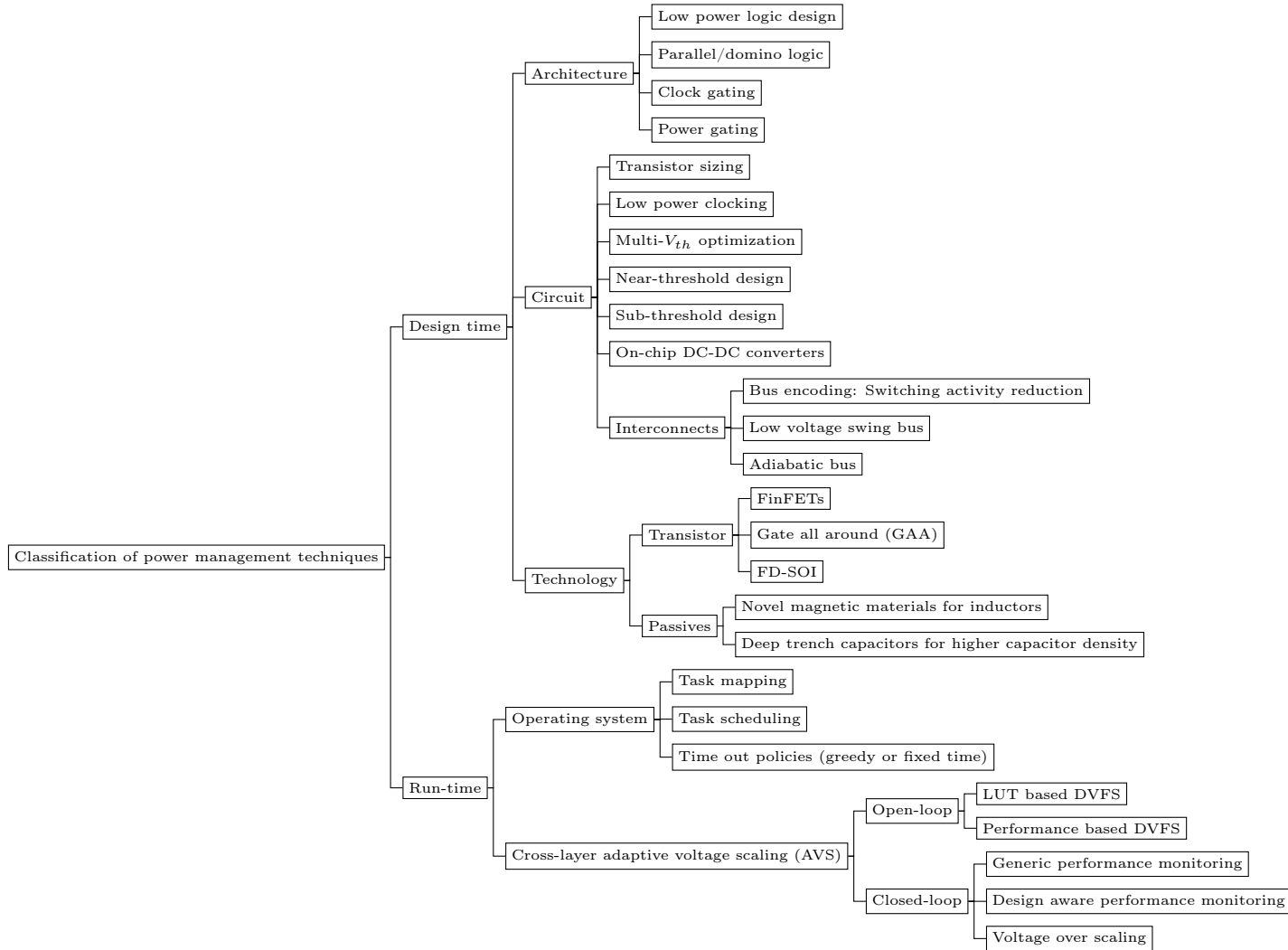


Figure 4.2: Taxonomy of power management techniques applied to chip multi-processor (CMP) systems.

4.4 System level power management

Almost all of the operating system level power management techniques require support at the hardware level (circuit or architecture implementations) [84]. Thread-to-core mapping and task scheduling [89] are made power aware to minimize the total dynamic power consumption of the chip multiprocessors (CMP). The decision to apply clock gating and/or power gating is typically made at the system level [84]. Modern general purpose processors support an extensive set of dynamic power states (P-states) and idle power states (C-states), which are power saving states implemented as a combination of dynamic operating voltages and clock frequency scaling in addition to clock and/or power gating. The power management techniques implemented at the operating system are applied during the execution of workloads on the CMP, and are, therefore, run-time power management techniques. As shown in Fig. 4.1, the primary parameters controlled by the operating system are the operating voltage and frequency, which are tuned either in an open loop or closed loop configuration. In open loop adaptive voltage scaling (AVS), a pre-characterized look-up table (LUT) is used to select an appropriate voltage and frequency pair to set the desired performance of a core for a given power budget. The AVS technique is classified as open loop as the hardware level variations due to process and the environment are not considered. In a closed-loop AVS system, the operating voltage and frequency of the cores are scaled after sensing the on-chip performance monitors such as the IPC (activity counters), operating temperature, critical path timing margin, and the dynamic and leakage currents. The performance monitors range from ring oscillators and latched tapped

delay lines to complex circuits characterizing the timing of the critical path(s). The decision to modify the operating voltage and frequency of a core is made locally by the LPM or globally by the GPM in a CMP. The GPM also considers the workload distribution across cores while scaling the voltage and operating frequency.

4.5 Cross-layer power management

A NSF funded workshop in 2012 [90] concluded that applying power management techniques in isolation at one of the four layers of abstraction: system, architecture, circuit, and technology, does not yield the optimum energy efficiency for a targeted performance and system reliability. One of the key findings of the workshop was that existing methods to define the power and energy efficiency of computing devices excluded the efficiency of the power conversion circuitry and the power distribution network (PDN). Other critical challenges identified at the workshop, which are considered in isolation for each layer in existing power management solutions, are

- statistical variability of circuit parameters attributed to process variation in sub-nanometer nodes,
- operating environment variability,
- workload uncertainty,
- deterioration in circuit performance due to device aging, and
- Absence of bi-directional information flow across the circuits, architecture, and system layers that enable cross-layer optimization.

The research described in this dissertation highlights the importance of accounting for the power conversion efficiency of the on-chip power delivery network including the DC-DC conversion circuits or the on-chip voltage regulators (OCVRs). A summary of OCVR circuit topology and system level considerations needed while designing the PDN with OCVRs is provided in Chapter 2. In this dissertation, the energy efficiency of a many-core system is modeled as a joint function of the power conversion efficiency of the OCVRs and the energy savings through dynamic power management techniques such as DVFS applied to the core [14,15,17]. The significant improvement in the energy efficiency of the many-core system, when the operating system is aware of the power conversion efficiency (PCE) of the OCVRs reinforces the need for cross layer power management.

4.5.1 Cross layer adaptive voltage scaling

Adaptive voltage scaling (AVS) schemes tune the supply voltage for a given performance (fixed operating frequency). Adaptive voltage and frequency scaling (AVFS) methods scale the operating voltage and frequency to optimize the dynamic power consumption. AVS and AVFS schemes are real-time closed-loop systems with on-chip circuits to sense and actuate the change in operating voltage and/or frequency. The real-time sensing of either one or several of the parameters such as the voltage [91], load current, junction temperature, and architecture level counters [92] are used in the feedback loop to drive the voltage and/or frequency adjustment circuitry (refer to Fig. 4.3). Unlike DVFS, where statistical models are used to assign a voltage level at a given frequency, AVS and AVFS methods identify the optimal power supply voltage

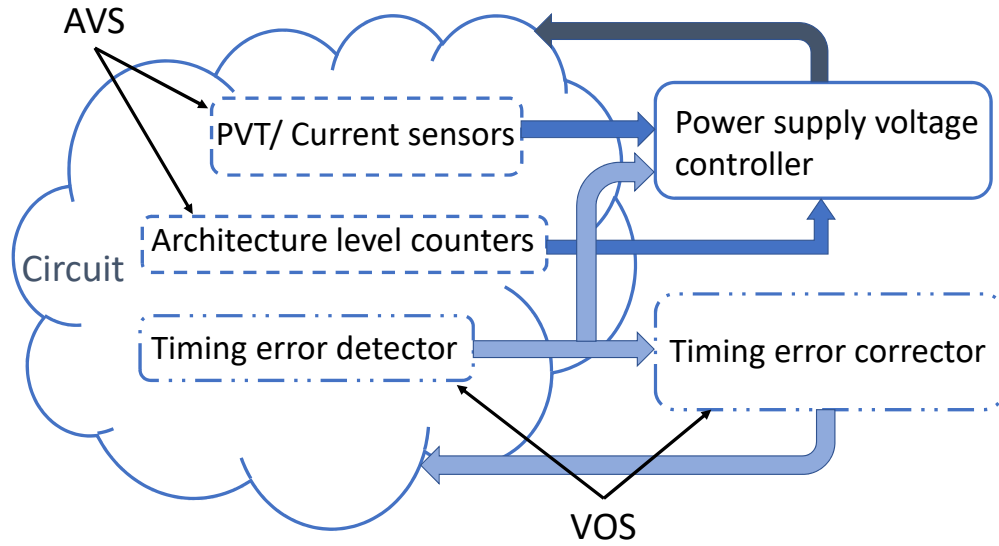


Figure 4.3: System model of a circuit implementing adaptive voltage scaling (AVS) or voltage over scaling (VOS).

at run-time.

A globally asynchronous and locally synchronous (GALS) system on a chip (SoC) is fabricated in 32 nm technology in [91]. The SoC includes a fine grained AVS system, where the appropriate supply voltage is determined by on-chip sensors that compensates for process, supply voltage, and temperature (PVT) variations. The PVT variations are sensed by distributed ring oscillators. Distributed timing fault sensors monitor the maximum functional frequency for a given supply voltage. The closed loop AVS scheme achieves up to 18.2% savings in energy consumption. The adaptive voltage generation technique does not require on-chip voltage regulators. Instead a V_{dd} hopping method is implemented to dither the supply voltage to each core at a pre-defined duty cycle, generating three distinct voltage levels.

4.5.2 Cross layer error-resilient voltage over scaling

Methods that scale the supply voltage such that timing errors are detected and corrected are classified as error-resilient voltage over scaling (VOS) schemes. Such methods are also implemented as a closed-loop to determine an optimal operating voltage (refer to Fig. 4.3). Error detection in critical paths is typically done through in-situ circuits in flip-flops [93]. Error resilient voltage techniques such as the Razor and RazorII [94] include an extra shadow latch to detect timing-errors and determine precise timing slack for margin shaving.

Error-resilient voltage scaling methods incur a throughput penalty when timing-errors are detected. In addition, a custom digital design flow for the insertion of error detection circuits is needed. As highlighted in Chapter 3, in sub-20 nm nodes, the voltage headroom for super-threshold voltage operation is low. The feasibility to implement a VOS technique is, therefore, limited without incurring a significant penalty in performance as over scaling the voltage results in near threshold or sub-threshold operation.

4.5.3 Thread to core mapping accounting for on-chip voltage regulator efficiency

A comprehensive body of research in cross-layer power management, completed in collaboration with George Mason University, is described in this subsection. The research explores thread-to-core mapping on heterogeneous CMPs that accounts for the power conversion efficiency of the OCVRs and a highly optimized point-of-load power delivery network to a core with run-time re-configurability of resources.

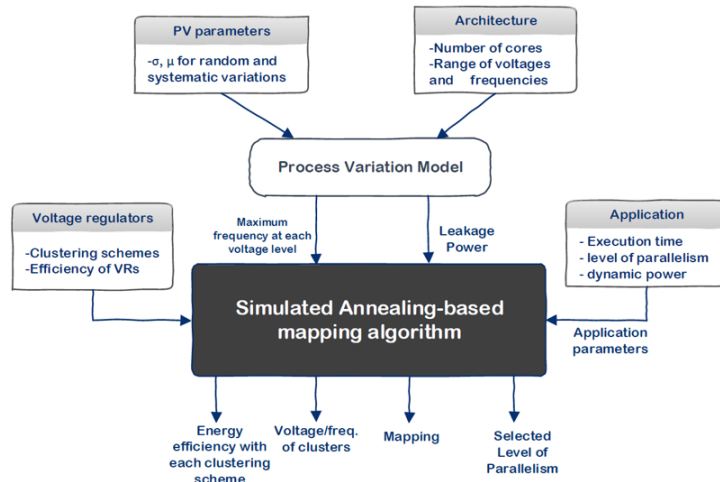


Figure 4.4: Simulation framework used to analyze the workload mapping algorithm in [95].

Per-core DVFS maximizes the savings in energy and meets the performance requirements of a given workload. Given a limited number of I/O pins and the need for finer control of the voltage and frequency settings per core, there is a substantial cost in using off-chip voltage regulators. Consequently, there has been an increased focus on the use of on-chip voltage regulators in many-core systems. However, integrating OCVRs comes at a cost of reduced power conversion efficiency (PCE) and increased complexity in the power delivery network and management of the OCVRs. The effect of the PCE of the OCVR on the thread-to-core mapping is investigated in [95]. The simulation framework to characterize the OCVR aware workload mapping algorithm is shown in Fig. 4.4.

The results from mapping threads using a simulated annealing technique indicate that up to 38% more energy is saved with PCE aware algorithms as compared to PCE-agnostic algorithms. The impact of the number of clustered cores and process

variation on the total energy efficiency of the system is analyzed. When relaxing the energy constraints by 10%, an effective thread-to-core mapping simplified the power delivery system by reducing the number of OCVRs required as compared to dedicating an OCVR per-core. The results indicate an important opportunity for system and circuit co-design to implement energy-efficient and complexity-effective platforms for a target workload.

4.5.4 Core scaling with enhanced power management

Heterogeneous architectures have emerged as a promising solution to improve the energy-efficiency of a system by allowing each application to run on a core that matches resource needs more closely than a one-size-fits-all approach. An ElasticCore platform where core resources along with the operating voltage and frequency settings are scaled to match the application behavior at run-time is proposed in [96]. A linear regression model for power and performance prediction is used to guide the scaling of the core size and the operating voltage and frequency to maximize the energy efficiency. Circuits that further optimize the power efficiency of the ElasticCore are developed. The ElasticCore with variable sizing of system resources and with dedicated point-of-load power delivery to each partition is shown in Fig. 4.5. The power delivery system of the ElasticCore is modeled using a configuration similar to the Intel Haswell processor [31]. Instead of serving the ElasticCore with a single OCVR that provides a high output current rating, each partition CW1 through CW4 is served with a dedicated OCVR with a maximum output current rating proportional to the peak current demand of the partition CW1. The current rating is defined

in terms of the maximum output current supplied by the voltage regulator. The PCE of a switching type OCVR is directly proportional to the load current and inversely proportional to the switching frequency [23]. Implementing multiple OCVRs of lower current rating, each serving a single partition of the ElasticCore, instead of a single OCVR with a high rating supplying current to all partitions provides multiple advantages:

- Variation in the power conversion efficiency is reduced as each OCVR supplies current to a load with smaller current variation.
- OCVRs with moderate peak current density and low switching frequency are used to serve each partition if a higher weight is assigned to energy efficiency rather than area efficiency.
- OCVRs with lower rating provide faster voltage transitions and, therefore, improved performance with DVFS.
- Reduction in the leakage power of the ElasticCore by selectively turning off the OCVRs serving unused partitions CW2 through CW4.

The results indicate that the ElasticCore is 30% more energy-efficient as compared to the ARM Big.Little architecture. In addition, the results of the OCVR aware system level scheduling technique indicate an important opportunity for system-architecture-circuit co-design for optimum energy management.

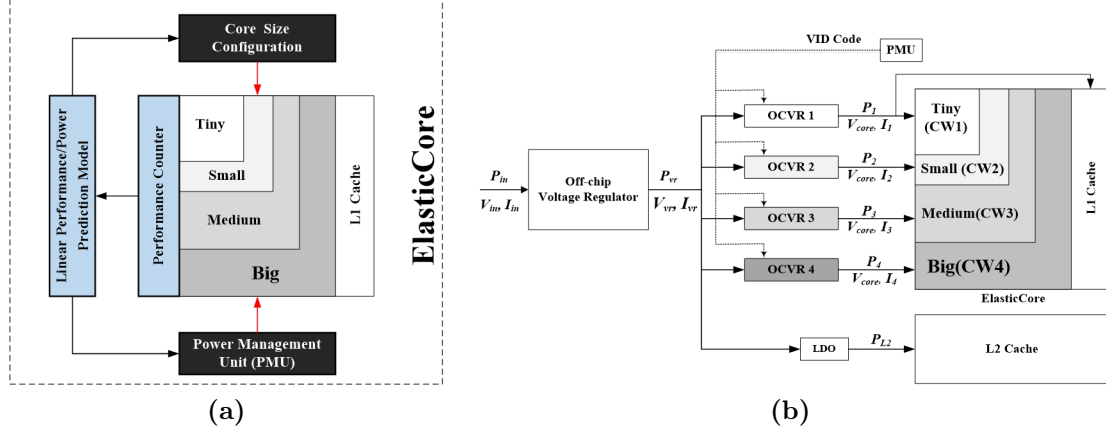


Figure 4.5: Representation of the Elasticcore (a) architecture and (b) two tiered power delivery configuration with a dedicated OCVR per partition [96].

4.6 Summary

In this chapter, existing methods for power management in an integrated circuit are discussed. The disparate methods operate at distinct layers of either the technology, circuit, architecture, or operating system. A need for cross layer power management is highlighted, specifically, the inclusion of the DC-DC converter circuits in the energy optimization problem of heterogeneous many-core systems. The power conversion efficiency of the voltage regulators integrated on-chip impacts the workload mapping and the energy efficiency of the system.

Chapter 5: Under-provisioned and reconfigurable power delivery

"The biggest problem with overdesign is knowing how much you've overdesigned."

– Oliver King, CTO, Moortec Semiconductor Ltd.

With the paradigm shift in computing systems from performance oriented design to energy efficient design, considerable research effort has focused on optimizing the core configuration by reducing the over-provisioning of the core resources. Little attention, however, is given to the reduction in the over-provisioning of the circuits delivering power to the cores. Conventionally, the voltage regulator and power conditioning circuits are off-chip. The power consumption and the footprint of the voltage regulators and the conditioning circuits is, therefore, not a concern while optimizing the power delivery to the core(s). The introduction of chip multi-processors (CMPs) resulted in new challenges in the delivery of power to the multiple cores. Providing low latency, dynamic voltage and frequency scaling (DVFS) per-core is challenging with off-chip voltage regulators [18]. The power supply voltage regulation is also reduced due to the longer on-chip interconnects connecting the off-chip VR to the multiple load circuits. On chip voltage regulators (OCVRs) have been extensively researched and successfully introduced in commercial multi-core systems in Intel 4th generation processors [31] as well as IBM POWER8 servers [30]. The choice of OCVR topology is dependent on several factors including system level parameters such as the

optimal power conversion efficiency and the maximum load current consumption, and the physical design of the passive components. System level tools such as a power virus [97] or McPAT [98] are used to determine a first order estimate of the peak power consumption of the cores, which is typically overestimated. As a result, the OCVR and the power delivery network are over-provisioned to support a peak load current larger than what is consumed by the cores.

In this chapter, an interconnected on-chip power distribution network is modeled. Rather than a static configuration designed for the worst case power consumption of the cores, a work load aware reconfigurable power delivery system is developed. A detailed statistical analysis of the cycle accurate power consumption profile of workloads executed on a CMP system is performed. Each OCVR is designed to support a peak current rating equal to the average load current I_{avg} consumed across all workloads. SPICE simulations indicate that by reducing the peak current rating of the OCVRs to support I_{avg} , the energy efficiency of the CMP improves and the on-chip area occupied by the OCVRs is reduced. A load balancing algorithm is developed for dynamic power management. The algorithm is executed on the on-chip power management unit (PMU), and is capable of reconfiguring the power delivery network (PDN) to combine the outputs of multiple OCVRs to support load currents in excess of I_{avg} .

The rest of the chapter is organized as follows: Prior work exploring reconfigurable power delivery networks (RPDN) is discussed in Section 5.1. The system level simulation and the power consumption profile of multi-application workloads are described in Section 5.2. The proposed power delivery methodology is discussed in Section

5.3. Simulated results verifying an improvement in the energy efficiency of the CMP system are also included in Section 5.3. Simulated results showing the improvement in the figures of merit of the OCVRs and the energy efficiency of the CMP system are provided, respectively, in Sections 5.4 and 5.5. SPICE simulation of a power grid benchmark to characterize the impact on power supply voltage droop during reconfiguration of the power delivery network is described in Section 5.6. A summary of the chapter is provided in Section 5.7.

5.1 Related work

Recent work has attempted to improve the energy efficiency of multi-core and many-core systems by reconfiguring the power delivery network dependent on the power demand of the work load. An RPDN using switched capacitor voltage regulators (SCVRs) and cross bar switches to serve eight cores is proposed in [99]. The RPDN consists of 32 cells, where each cell is an SCVR capable of supporting two voltage step down conversions (2:1 and 3:2). The simulation results indicate that the reconfigurable power delivery network offers 40% energy savings as compared to a configuration with per core voltage regulation. The switched capacitor voltage regulators offer a power conversion efficiency (PCE) of 80%. The work does not address the inferior voltage regulation offered by the SCVRs. The SCVR topology is selected over a buck converter as scaling the SCVR, which involves reducing the size of the capacitor and MOS switch for smaller load currents with moderate losses is less complex. Multi-phase buck converters offer the same advantage with superior voltage regulation. A fully integrated voltage regulator (FIVR) [31], based on a buck

converter topology offers a PCE of 90% and, therefore, better energy efficiency. The SCVR offers a voltage transition time in the range of 1000 ns. The voltage transition time for the FIVR is 500 ns (0 V to 1 V). The selection of the SCVR is, therefore, not optimal as OCVRs with improved PCE, output voltage regulation, and voltage transition time are available. The area, power, and transient time of the switching network in the RPDN is also not analyzed. The scalability of the power network is achieved by partitioning the RPDN to serve clusters of cores. With a minimum of four cores per cluster (to support the four voltage levels of the cluster), the number of switches required in the RPDN scales as $N \times M^2$, where M is the number of cores in a cluster and N is the number of clusters in the system. Each cluster requires a dedicated power grid. The area and power overhead due to the switching network and the power grid is not analyzed for a many core platform.

A run time reconfigurable voltage regulator (VR) network of buck converters is described in [100]. The lowest energy consumption across various DVFS levels is determined by solving an integer linear programming (ILP) problem. The timing penalty to set the switching network is not quantified, and the ILP is solved for discrete DVFS timing penalties ranging from 5% to 15%. An off-chip buck converter (LTC3816) SPICE model is used instead of an OCVR, although the OCVR offers an order of magnitude faster voltage response time when implementing DVFS [18].

In [101], an on-chip RPDN that includes dynamic voltage and frequency scaling is proposed on a two tier 3-D IC. The cores are grouped according to voltage demand, which is determined from a look-up-table comprised of the power envelopes (maximum power consumption per time slot) per control cycle and the associated voltage levels

obtained from tracking and predicting the power signature using an autoregression algorithm. The power signatures for SPEC CPU2000 benchmarks are obtained through simulations using *Wattch*. Extraction of the power phase is completed through singular value decomposition of the covariance matrix of the power signatures per time slot and per workload.

A workload scheduler is developed to minimize the power slack (the difference between a predetermined power threshold and power envelop per time slot). The penalty of MIPS/Watt due to time multiplexing of the connections between the core and VRs is not analyzed in [101]. In addition, an underlying assumption of the workload scheduling algorithm is that in any given time slot, there is always a subgroup of cores available with positive power slack, which does not always hold true. Since the maximum number of cores that are served by a single VR is limited, if the total driving capability of the system for all VRs is less than the total number of cores, no plausible connections between cores and VRs exists that meets the total current demand. In addition, the network switching time is three times the time slot considered by the power management algorithm, which significantly impacts the performance when switching the workload from one subgroup to another. The power loss in the VRs and switching network is not accounted for when determining the power savings of the RPDN proposed in [101].

Clustering of VRs to boost the energy efficiency of the system is proposed in [23, 102], but the work ignores the variation in the power conversion efficiency of the low dropout (LDO) VRs due to dynamic voltage and frequency scaling. As shown in [95], ignoring the variation in the PCE of the VRs leads to suboptimal workload

mapping and, therefore, a large penalty on the energy savings possible with DVFS.

Recent work on RPDNs does not provide an analysis of the penalty in the response time of the power delivery system due to the search and decision time needed to flip the requisite number of switches and reconfigure the connections between the cores. In addition, all prior work consider over-provisioned VRs designed for worst case power consumption. The RPDN developed in [99–101] improve the energy efficiency of multi-core systems through the implementation of a dynamic power distribution network. The reconfiguration of the PDN is achieved through a CMOS switch network controlled by a power management unit. The three publications include different power converters for implementation of the RPDN. While [100] considers a commercial off-chip buck converter model, an on-chip SIMO converter is considered in [101] and a switched capacitor OCVR is considered in [99]. The power and area overheads of the RPDN due to the OCVR and switching network are well analyzed at the circuit level. The reconfiguration of the RPDN either in space or time is dependent on accurate prediction of the power profile of the workloads executed on the cores. Different heuristics are used in the three publications including off-line lookup tables that store the power/voltage information and an auto regression algorithm to predict the power signatures at run time. The main techniques used in the three publications are summarized in Table 5.2.

5.2 Power dissipation in CMPs

The behavior of each application differs with regard to the utilized resources of a computing system. While some applications are cpu intensive, others are memory

Table 5.1: Architectural parameters of the core.

Parameter	Value
Core clock frequency	2.4 GHz
Power supply voltage (V_{dd})	1 V
Issue, commit width	2
INT and FP instruction queue	16 entries
Load and Store Queue	16 entries
INT and FP Physical Register File	48 entries
ROB size	48
L1 cache	32KB, 4-way
L2 cache	256KB

intensive. The power dissipation pattern, therefore, varies across different applications. In addition, applications exhibit differing power dissipation behavior in different execution phases. The power dissipation pattern of workloads must, therefore, be accounted for while designing the power delivery system. The analysis of the typical power consumption profile of different workloads offers insight on the circuit level implementation of the OCVRs that provide regulated power to the CMP system. The simulation methodology to determine the power consumption of SPEC CPU benchmarks and the corresponding simulated results are described, respectively, in Sections 5.2.1 and 5.2.2.

Table 5.2: Summary of important system, architecture and circuit parameters used to design the RPDNs in [99–101].

Publication	Implemented VR	Power prediction algorithm	Objective function for ILP	Algorithm/ILP executed off-line or at Run-time	CMP system configuration	Peak PCE	VRs to cores ratio	Max. system power savings	Max. system energy savings	Benchmark for which max. power/energy savings are obtained
[100]	Off-chip buck converter	None	Minimize power across DVFS levels for a given time interval s.t. workload execution delay is below a certain DVFS timing penalty	Off-line	16 cores CMP modeled on Intel Nehalem architecture. Technology node not specified.	80%	1:1	Not given	9%	Barnes
[101]	SIMO buck converter	Auto Regression	Minimize number of voltage levels	Off-line and run-time	64 core, 130 nm	77%	1:3	43%	Not given	All SPEC2000 benchmarks
[99]	SCVR	Booster sync+ [103] and hint instructions	None	Run-time	8-core, 65 nm	81.8%	4:1	40%	40%	pbbs-mm

5.2.1 Simulation methodology

A 16-core CMP in a 45 nm technology is modeled using a processor architectural simulator [104]. McPAT [98] is integrated in the simulator to analyze the power consumption of the core. Each core has a 2-way issue and out-of-order execution unit. The micro-architectural parameters of the core used in simulations are summarized in Table 5.1.

A set of 49 applications from the SPEC CPU2000 and SPEC CPU2006 benchmark suites are studied to determine the power dissipation behavior of the core. Each benchmark is simulated at four timing intervals to analyze execution phases with differences in power consumption profiles. The simulations are run for 10K cycles per time interval, and the power consumption is sampled cycle by cycle.

5.2.2 Analysis of the power dissipation behavior

The statistical variation of the power consumption per cycle of different SPEC CPU benchmarks with single phase forwarding is shown through a box plot in Fig. 5.2. The interquartile range for all the studied benchmarks falls approximately an order of magnitude below the peak power consumption of 5.73 W reported through McPAT. The number of outliers beyond 5σ coverage for each benchmark is an insignificant fraction of the sample size.

The consolidated power consumption and power variation histograms of the studied benchmarks are shown in Fig. 5.1. The dashed line in Fig. 5.1(a) delineates the average power consumption across all benchmarks, which is approximately 0.55 W. The power dissipation of the applications is between 0.3 W and 0.5 W for approxi-

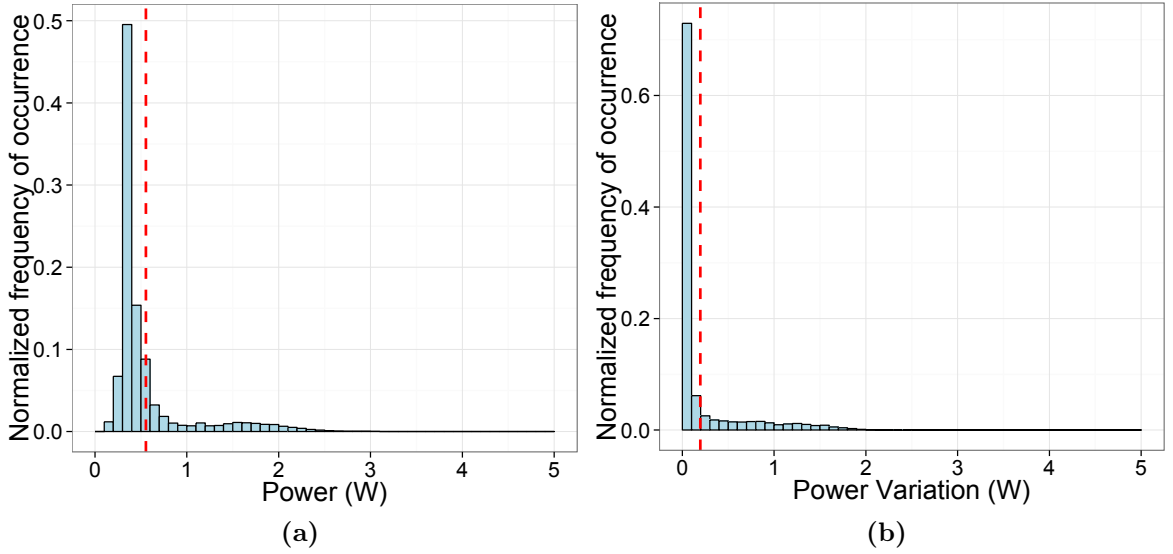


Figure 5.1: Histogram of (a) power dissipation, and (b) power variation per cycle, for the 49 SPEC CPU2000 and SPEC CPU2006 benchmarks.

Table 5.3: Combined power dissipation characteristics of SPEC CPU2000 and SPEC CPU2006 benchmarks.

Parameter	Minimum	Average	Maximum
Power dissipation (W)	0.175	0.555	4.755
Power dissipation variation (W)	0	0.195	4.333

mately 65% of the execution time. The studied benchmarks consume less than the average power for more than 78% of the run-time. The variation in power dissipation between any two clock cycles averages 0.2 W, as shown in Fig. 5.1(b). The power variation is less than 0.1 W for about 90% of the run-time. The characterization of the power for the SPEC CPU2000 and SPEC CPU2006 benchmarks is summarized in Table 5.3. The peak power P_{peak} of 5.73 W reported by McPAT is never consumed. The maximum power consumption of 4.75 W is consumed for a small percentage ($7.5 \times 10^{-5}\%$) of the run-time across all workloads.

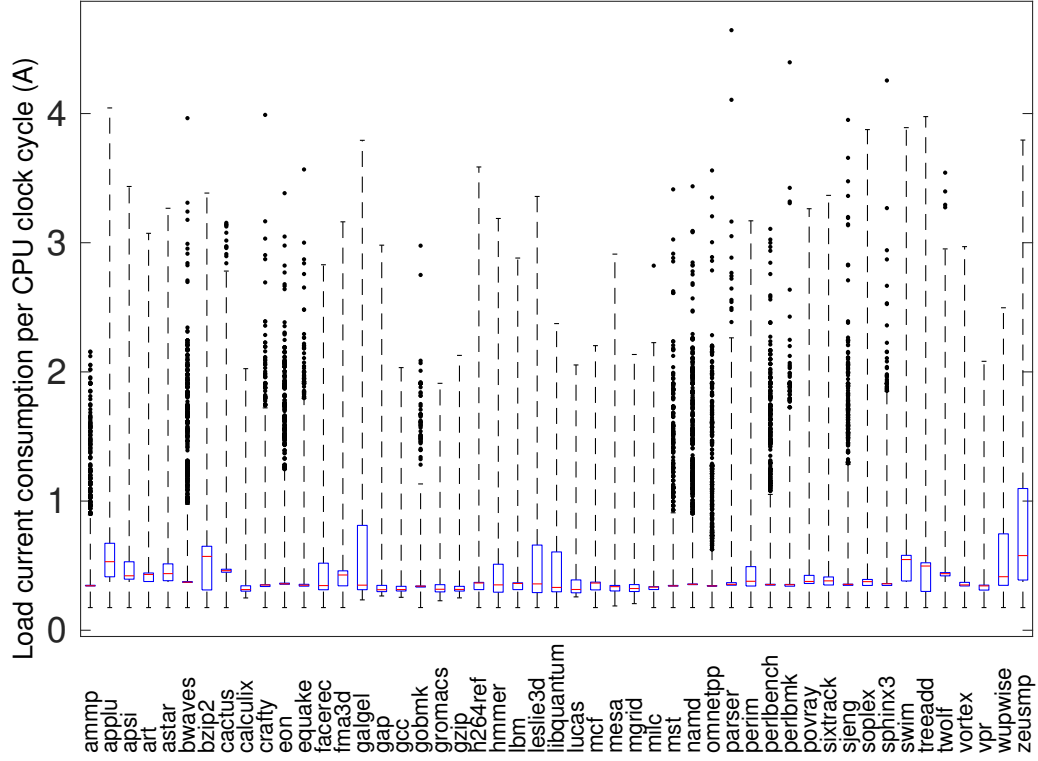


Figure 5.2: Statistical analysis of per cycle power consumption of SPEC CPU benchmarks.

5.3 Proposed interconnected power delivery network

The power consumption characteristics listed in Table 5.3 indicate that the VR rating is over-provisioned for the majority of the run-time of the workloads. Significant work has been done to optimize the core configuration, work load mapping, and dynamic/static clustering of the cores served by the off-chip VRs, but the energy and area loss incurred due to the integration of over-provisioned VRs has been overlooked. In the proposed power delivery network, on-chip voltage regulators with a maximum output current equal to the average current consumption I_{avg} of the load circuits (cores) are considered. A technique to deliver currents higher than I_{avg} is described

in this section.

The block representation of the proposed interconnected power delivery network is shown in Fig. 5.3. For a CMP system consisting of N cores, N OCVRs provide the regulated power. As shown in [95], providing a single OCVR to an optimally sized cluster of cores yields similar benefits in energy savings as per core DVFS, while also reducing the number of OCVRs as compared to the number of cores. In the proposed PDN, if the OCVR serves a cluster of n cores, the power rating of the OCVR is $\sum_{i=1}^n I_{avg,i}$. The output of each OCVR is connected to the inputs of a high speed switching (HSS) fabric. The N outputs of the HSS fabric are connected to the local PDN grid of the N cores or core clusters. The HSS fabric is controlled by the power management unit (PMU). The interconnected power delivery network shown in Fig. 5.3 provides increased service reliability as compared to the conventional radial topology used for on-chip power distribution [23]. In addition, the interconnected network provides opportunity to balance the load current through reconfiguration of the switches. The current sensors placed in each core are constantly monitored by the PMU. When the sum of the currents sensed from all cores within a cluster (I_{sense}) reaches a threshold ΔI below I_{avg} , the PMU configures the HSS to source additional current from the OCVRs that are operating at the same power supply voltage level under DVS controlled by the PMU.

The high speed switches are controlled by logic within the PMU that operate on two system parameters, the V_{dd} levels and the total load current sensed from each core cluster. The analysis of the power consumption of the workloads provided in Section 5.2 indicates that the probability of the load current exceeding I_{avg} is 22%.

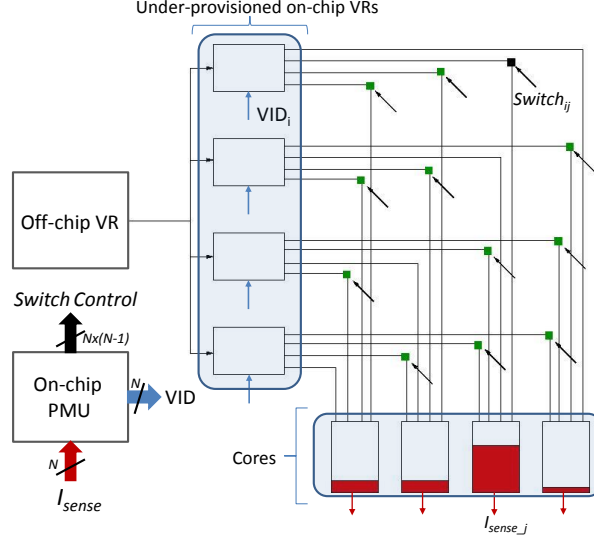


Figure 5.3: Proposed interconnected on-chip power delivery network with run-time voltage regulator clustering through a switching fabric.

As a result, there are always more than one core clusters operating at or below I_{avg} . The PMU is provisioned to add at least one additional OCVR to serve a cluster or core requiring current higher than I_{avg} . An available OCVR is ensured if the number of DVFS levels is less than the number of core clusters in the CMP system. The sum of the decision time of the PMU and the time to reconfigure the switches must be less than or equal to the load current transient response time (current slew rate) of an OCVR with a current rating of I_{peak} to ensure an uninterrupted power supply to the core or cluster of cores. The switching control of the HSS fabric is described by *Algorithm 5.3a*.

Algorithm 5.3a is executed on the on-chip PMU. The inputs to the PMU are the current power supply voltage and load currents sensed for each core served by a dedicated OCVR. The PMU controls the $N \times (N-1)$ HSS matrix ($Switch_{N,N-1}$)

to dynamically configure the connections between the N OCVRs and N cores. Two dynamic arrays, $CORE_RED$ and $CORE_GREEN$, are maintained with the identification label of the cores, which are consuming, respectively, more than I_{avg} and less than $I_{avg} - (\Delta I)$ currents. If the length of the array $CORE_RED$ is non zero, the $OCVR_CLUSTER$ routine is executed, which combines the outputs of the OCVRs with total output current less than two times I_{avg} , which is the maximum output current of the OCVR for a given core with average load current consumption of I_{avg} . The output voltage of the combined OCVRs is matched before activating the corresponding switches in the HSS matrix. If the dynamic array $CORE_RED$ is empty, the routine $OCVR_DECLUSTER$ is executed, which deactivates the switches that combined the outputs of the OCVRs. The two constraints in the execution of *Algorithm 5.3a* are 1) the time taken to reconfigure the connections between the OCVRs and cores is less than one core clock cycle ($t_{switch} + t_{PMU} < t_{core}$) and 2) the combined power consumption of all the N cores in the system ($\sum_{i=1}^n V_x \cdot I_{sense_x}$) is less than the total power delivered by the N OCVRs ($N \cdot V_{dd_m} \cdot I_{avg}$). The workload scheduling heuristic described in Chapter 6 assures the constraints are met.

Algorithm 5.3a is implemented in the Python programming language and is analyzed with the parameters summarized in Table 5.4. The per cycle power consumption of different SPEC benchmarks for a finite number of CPU cycles is used as an input to *algorithm 5.3a*. In addition, a stochastic model of the current consumption of the cores in a CMP system is developed based on the statistical parameters captured from the per cycle power consumption analysis of the SPEC CPU benchmarks (see Fig. 5.1). The dynamic power consumption $P_{dynamic}$ per core is modeled as a Type IV

Algorithm 5.3a Load balanced power delivery with run-time OCVR clustering to support higher than average load current consumption.

Inputs:

Current consumption sensed from the core: I_{sense_x}

Current threshold: ΔI

Voltage level applied to the core: $V_x \in [V_{dd_1}, V_{dd_2}, \dots, V_{dd_m}]$, where $x \in [1, \dots, N]$,
 m = number of DVFS levels, N = number of OCVRs/Cores

Switch matrix: $Switch_{N,N-1}$

Constraints:

$$t_{switch} + t_{PMU} < t_{core}$$

$$\sum_{i=1}^n V_x \cdot I_{sense_x} < N \cdot V_{dd_m} \cdot I_{avg}$$

- 1: Append array *CORE_RED* with core id x where $I_{sense_x} \geq I_{avg} - \Delta I$
 - 2: Append array *CORE_GREEN* with core id y where $I_{sense_y} < I_{avg} - \Delta I$
 - 3: **if** $length(CORE_RED) > 0$ **then**
 - 4: call *OCVR_CLUSTER* $\triangleright$ Reconfigure the PDN by clustering the output of OCVRs
 - 5: **else**
 - 6: call *OCVR_DECLUSTER* $\triangleright$ Reconfigure the PDN by de-clustering the output of OCVRs
 - 7: **procedure** *OCVR_CLUSTER*
 - 8: **for** each i in $\min(length(CORE_RED), length(CORE_GREEN))$ **do**
 - 9: $Demand(i) \leftarrow CORE_RED(I_{sense_i}) + CORE_GREEN(I_{sense_i})$
 - 10: **if** $Demand(i) \leq 2 \cdot I_{avg}$ **then**
 - 11: $V_{CORE_RED(i)} \leftarrow V_{CORE_GREEN(i)}$ $\triangleright$ Align the V_{dd} levels of the two cores whose OCVR outputs are being combined
 - 12: $Switch_{CORE_RED(i), CORE_GREEN(i)} \leftarrow 1$ $\triangleright$ Close the switch so that $CORE_RED(i)$ is served by the OCVR connected to $CORE_GREEN(i)$
 - 13: Delete $CORE_RED(i)$ and $CORE_GREEN(i)$ from the respective arrays
 - 14: **procedure** *OCVR_DECLUSTER*
 - 15: **for** each nonzero element in sparse matrix *Switch* **do**
 - 16: **if** $I_{sense_i} + I_{sense_j} < 2 \cdot I_{avg} - \Delta I$ **then**
 - 17: $Switch_{i,j} \leftarrow 0$ $\triangleright$ Open the switch connecting core i with OCVR j
-

Pearson distribution [105] given by (5.1). The parameters m , a , ν , and λ are derived from the skewness and kurtosis exhibited by the per cycle power consumption of the SPEC CPU2000 and SPEC CPU2006 benchmarks and are, respectively, 4.145, 2.277, 0.889, and 0.322. The load current consumption as obtained from the stochastic model for 1,000 CPU cycles across 16 cores is shown in Fig. 5.4. The peak current rating of each OCVR is set to 0.6 A, which is one order of magnitude less than the I_{peak} obtained through McPAT. The high speed switch configuration to support the run-time load current variation on each core for two randomly chosen time stamps is shown in Fig. 5.5. A case where the current consumption of core 7 exceeds 2.5 times the maximum current output of the OCVR is shown in Fig. 5.5(c). The algorithm clusters the output of the OCVRs available in the CMP that are connected to cores demanding less than the maximum current output of a single OCVR. A statistical load current model with normal distribution is also analyzed through Monte Carlo simulations with a maximum possible value of I_{peak} . The frequency of OCVR clustering increases with a normally distributed load current model, however, a developed workload mapping technique, described in Chapter 6, significantly reduces the occurrence of clustering. The four V_{DD} levels listed in Table 5.4 are selected corresponding to the core configuration provided in Table 5.1.

$$P_{dynamic} = f(x)dx = \left[1 + \left(\frac{x - \lambda}{a}\right)^2\right]^{-m} \exp\left[-\nu \cdot \tan^{-1}\left(\frac{x - \lambda}{a}\right)\right] dx \quad (5.1)$$

The RPDN described in this chapter is novel, as the PDN is designed to sup-

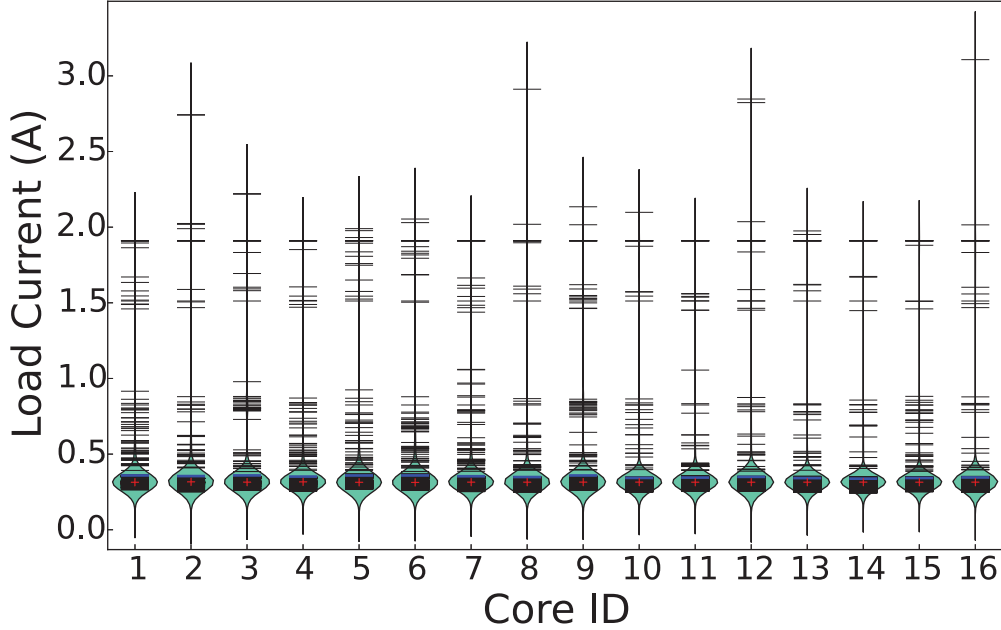


Figure 5.4: Load current variation of a 16-core CMP system with 16 OCVRs each with a peak rating of 0.6 A. The stochastic model for load current consumption across the cores is based on SPEC CPU2000 and SPEC CPU2006 benchmark power trace analysis. The variation in load current is used as an input to *Algorithm 5.3a*.

ply the typical power demand of the cores with the capability to support the peak power demand if necessary. The PDN configuration is managed dynamically without the overhead of solving any off-line or on-line linear or nonlinear programming optimization problem.

5.4 Improvement in figures of merit of OCVRs and other system parameters

The increasing power density of CMPs requires efficient OCVRs with a small area and a fast slew rate. The figures of merit of the OCVR depend on the circuit topology and the target operating point, the regulated output voltage at the maximum

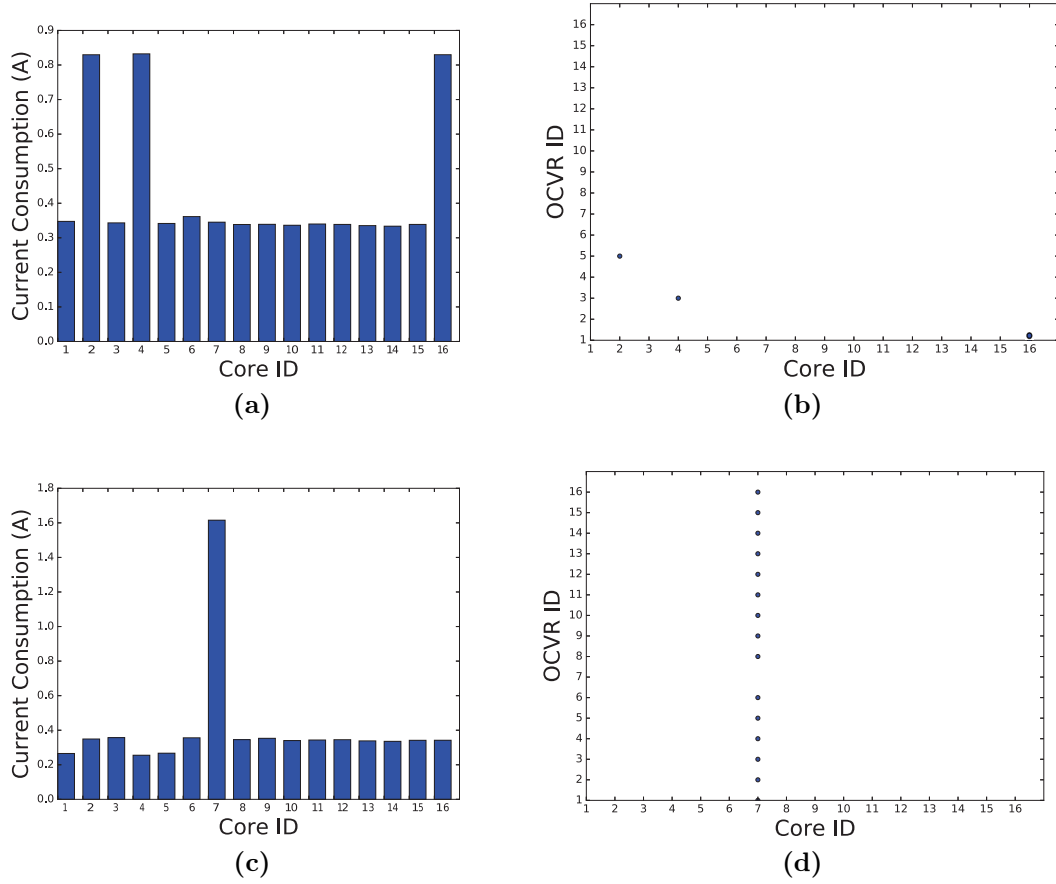


Figure 5.5: Simulated results of the implementation of Algorithm 5.3a. A 16-core CMP with 16 OCVRs each with a peak rating of 0.6 A is considered. The load current across cores is shown for two time stamps in (a) and (c) with the corresponding switches that are active to supply current that exceeds 0.6 A shown in, respectively, (b) and (d).

output current. The choice of DC-DC voltage converter topologies includes buck converters, low drop-out (LDO) voltage regulators, or switched capacitor voltage regulators (SCVR). The LDO offers a high power conversion efficiency (PCE) with the smallest foot-print. However, the PCE of the LDO varies linearly with the V_{out}/V_{in} ratio [24, 30]. Considerable power is, therefore, lost in the LDO at lower operating voltages when applying DVFS. The SCVRs are capable of stepping the input voltage

Table 5.4: Simulation parameters for *Algorithm 5.3a*.

Parameter	Value
Number of cores	8, 16, 32, 64, 128
V_{DD} levels	0.7 V, 0.8 V, 0.9 V, 1 V
OCVR peak current rating I_{avg}	0.6 A
Current threshold (ΔI)	0.1 A
Duration of execution	10 million CPU cycles
Load current variation	Stochastic model with Pearson type IV distribution based on SPEC CPU benchmarks

up or down based on the configuration of switches and capacitors. The power loss in the resistive switches due to conduction and frequent switching limit the use of SCVRs in on-chip voltage conversion. The output voltage regulation of the SCVR topology is inferior to other VR topologies due to 1) the strong output voltage dependence on the current demand of the load circuit and 2) the feedback for output voltage regulation is difficult to implement. The output voltage regulation is improved by increasing the switching frequency of the MOS switches, which requires a wider width and therefore results in higher dynamic power consumption, producing a lower PCE. Due to an inferior voltage regulation and PCE, switched capacitor voltage converters are not considered in this chapter. The switching DC-DC buck converter offers superior power supply voltage regulation and a wider output voltage range with a comparable PCE to LDOs. A buck converter consists of a switching network and a passive low pass filter. The inductor in the low pass filter acts as a low-loss energy transfer device that improves the power conversion efficiency. Buck converters are, therefore, an ideal choice to power processing elements that implement DVFS. After analyzing each OCVR topology and accounting for the proposed RPDN, the buck converter is

Table 5.5: Improvement in the figures of merit of the OCVR supporting I_{avg} instead of I_{peak} [24].

Area reduction	$0.16 \times Area$
PCE	$1.3 \times PCE$
Output voltage ripple reduction	$0.1 \times V_{ripple}$
Improvement in load regulation	$10 \times \beta_{load}$
Line regulation	$1 \times \beta_{line}$
Reduction in startup time	$0.5 \times T_{startup}$

selected for integration with the CMP system considered in this chapter.

5.4.1 Improvement in figures of merit of a buck converter

The figures of merit of a buck converter are analyzed for changes to the peak load current rating of the regulator. The goal is to characterize the impact on the OCVR response time and energy efficiency when designing the OCVRs to support only the average load current demand of the cores. The switching DC-DC buck converters with an input voltage of 2.5 V and an output voltage of 1 V are simulated using TI Webench [106]. Buck converter configurations are simulated with varying maximum output current or peak current rating.

The improvement in the figures of merit of a buck converter with peak current rating of I_{avg} as compared to I_{peak} are summarized in Table 5.5. The ability to maintain a constant output voltage with changes in the load current is characterized by β_{load} and with changes in the input voltage by β_{line} [24]. The time taken to stabilize the output voltage to the desired regulated value is given by $T_{startup}$. The startup time of the buck converter for a peak rating of I_{avg} and I_{peak} [106] is shown in Fig. 5.6.

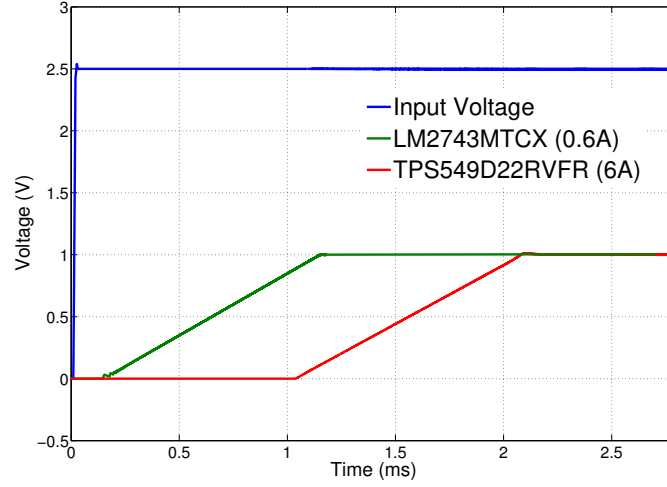


Figure 5.6: Startup time of the voltage regulators designed with a peak load current rating of 6A (I_{peak}) and 0.6A (I_{avg}).

5.4.2 Improvement in power conversion efficiency of a buck converter

The power consumed by the buck converter P_{buck} is given by (5.2) [24]. The P_{mos} , P_{ind} , P_{cap} , and P_{pwm} are the power loss in, respectively, the MOS power transistors and the cascaded buffers driving them, the inductor of the filter circuit, the capacitor of the filter circuit, and the pulse width modulator circuit. The detailed mathematical formulae of each of the components that contribute to P_{buck} are given in [24], [107].

$$P_{buck} = P_{mos} + P_{ind} + P_{cap} + P_{pwm} \quad (5.2)$$

The power consumed by the filter circuit, power transistors, and the buffers driving the power transistors increases with the maximum supported output current of the buck converter. Alternatively, multiple phases are used to drive higher output

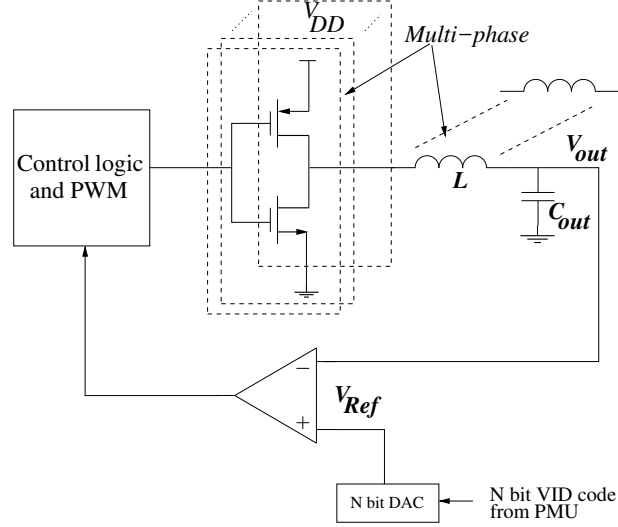


Figure 5.7: Schematic of a multi-phase DC-DC switching buck converter.

currents. The circuit schematic of a buck converter with multiple phases of the filter circuit, MOS power transistors, and cascaded buffers is shown in Fig. 5.7.

Two custom buck converters with a maximum output current rating of 6 A and 0.6 A are implemented [106]. The two converters represent voltage regulators that support the I_{peak} and I_{avg} currents of a CMP with core parameters listed in Table 5.1. The size of the filter inductor is chosen such that the percentage of peak current ripple $I_{pp,L}$ remains the same even with an order of magnitude reduction in the peak current drive. The theoretical calculations of an approximately one-third reduction in P_{buck} when using the smaller converter are verified through simulation. The power consumption of the various components of the buck converter along with the occupied area is listed in Table 5.6. The on-chip implementation of the two buck converters yields similar ratios between the power consumed by each component, although at a higher switching frequency for the smaller regulator to reduce the size of the filter

inductor and capacitor.

The large reduction in the power dissipation of the buck converter due to a reduction in the peak load current rating results in an improvement in the power conversion efficiency (PCE). Although the over-provisioned buck converter offers the same peak PCE at an output current of 6 A as the under-provisioned buck converter at a current of 0.6 A, the reduction in the PCE with decreasing output current for the over-provisioned converter is significant (see Fig. 5.8). The typical workloads executed on a CMP with the core configuration listed in Table 5.1 consume currents less than I_{avg} for 70% of the execution time. The buck converter with an output rating of 0.6 A, therefore, offers a higher average PCE for a majority of the run-time of the workloads. Circuit techniques like automatic mode switching between pulse frequency modulation (PFM) and pulse width modulation (PWM) are used to boost the light load efficiency of buck converters. However, PFM mode in buck converters suffers from inferior transient response to changes in load current as compared to PWM [14]. In addition, the output voltage ripple is greater in PFM mode as compared to PWM mode. The limited transient response and large output voltage ripple deteriorate the line and load regulation offered by the OCVR. Despite the disadvantages of enhancing the light load efficiency of buck converters, in Section 5.5, buck converters with enhanced light load efficiency are simulated to analyze the impact on the energy efficiency of a CMP system.

Table 5.6: Power consumption of DC-DC switching buck converters designed for 0.6 A and 6 A peak load currents.

Maximum load current	$I_{pp,L}$	Switching frequency	PWM duty cycle	P_{mos}	P_{ind}	$P_{cap}+P_{pwm}$	P_{buck}	Foot-print
6A	2.4A	695 KHz	20.90%	319.97 mW	315 mW	242.13 mW	889.71 mW	273 mm ²
0.6A	0.24A	3 MHz	26.84%	197.91 mW	43.56 mW	58.97 mW	300.44 mW	63 mm ²

Table 5.7: Properties of the switching DC-DC buck converters simulated for maximum power conversion efficiency [106].

Device	LM2743MTCX	TPS548A20RVER	TPS548A20RVER	TPS548A20RVER	LM2743MTCX	TPS549D22RVFR
Maximum output current (A)	0.6	2	3	4	5	6
Efficiency at maximum output current	96.36	94.83	91.3	96.6	94.6	94.44
High efficiency at light load feature	No	No	No	Yes	Yes	Yes
Peak-to-peak inductor ripple current (A)	0.18	0.55	0.79	1.26	0.83	1.22
Switching frequency (kHz)	100	234.2	231.1	100	328.13	327.5
Duty cycle (%)	40.16	40.64	42.6	40.3	40.77	41
Peak-to-peak output ripple voltage (mV)	2.28	1.83	2.69	6.3	1.15	1.69
Total power dissipation (mW)	22.65	109.1	285.3	141.3	285.85	353.24

5.5 Energy efficiency of a CMP with RPDN and under-provisioned voltage regulators

The total energy consumption of a CMP implemented with a conventional power distribution network (PDN) for a given execution time T_{epoch} with N cores and N OCVRs is given by (5.3). The cores are served by over-provisioned OCVRs identical to the buck converter with a maximum output current of 6 A. The dynamic and static power consumed by cores that implement DVFS are given by $P_{dynamic}$ and P_{static} , respectively. PCE_1 represents the power conversion efficiency of the over-provisioned OCVR. At low load currents close to I_{avg} , the PCE_1 offered by the over-provisioned buck converter is 80%. Alternatively, if the power delivery system is designed with each core supported by a buck converter that supplies a maximum output current of 0.6 A, the achieved PCE_2 at I_{avg} is 96.36%. In addition, the static power consumed by the cores or core clusters is zero as the idle core(s) are power gated through the high speed switching (HSS) fabric. The HSS fabric, however, imposes an additional switching loss P_{switch} , which is the dynamic power consumed by the PMOS transistors while switching, and a conduction loss $P_{conduction}$ while in the ON state and passing the average current I_{avg} .

$$E_{CMP,conventional} = \left\{ \sum_{i=1}^N \frac{(P_{dynamic_i} + P_{static_i})}{PCE_1} \right\} \cdot T_{epoch} \quad (5.3)$$

The total energy consumed by the CMP with N OCVRs, where each OCVR is designed for an I_{avg} rating, and $N \times (N-1)$ PMOS switches is given by (5.4). The parameters j , k , and l are, respectively, the number of active core(s) consuming current

below I_{avg} , the number of active core(s) consuming current above I_{avg} , and the number of idle core(s) power gated through the HSS network. In the case of idle cores, the power consumed by the OCVRs ($I_{quiescent} \cdot V_{out}$) is the only component contributing to the system energy. As described in Section 5.2, applications consume current less than I_{avg} for approximately 78% of the run-time. The P_{switch} loss is incurred for 22% of the execution time of the workloads when the load current exceeds I_{avg} .

$$\begin{aligned}
 E_{CMP,proposed} = & \sum_{t=1}^{T_{epoch}} \left\{ \sum_{i=1}^j \frac{(P_{dynamic_i} + P_{static})}{PCE_2} \right. \\
 & + \sum_{i=1}^k \frac{(P_{dynamic_i} + P_{static} + P_{switch} + P_{conduction})}{PCE_2} \\
 & \left. + \sum_{i=1}^l I_{quiescent} \cdot V_{out} \right\}; \\
 & j + k + l = N
 \end{aligned} \tag{5.4}$$

Circuit simulation of a PDN designed to support I_{avg} for each of 16 cores is performed to characterize the energy consumption of a CMP, with results compared to (5.4). The 16 cores are simulated as piecewise constant current sinks. The current variation for the 16 sinks is shown in Fig. 5.4 for 1,000 consecutive clock cycles. A 16x15 PMOS switching network is implemented in a 45 nm technology. The gates of the PMOS switches are controlled through time varying voltage signals. The P_{switch} and $P_{conduction}$ for a PMOS switch with an output capacitance provided by a single core is determined through SPICE simulations. The P_{static} and $P_{dynamic}$ of a core is measured through McPAT. The $P_{dynamic}$ for each core is overestimated as the power

Table 5.8: Parameters of the PDN determined through SPICE simulation.

Parameter	Value
PMOS width	800 μm
PMOS switching time	160 ps
Area occupied by 16x15 switching network	9600 μm^2
Per core maximum switching capacitance	2.4 nF
Piecewise constant load current variation	Stochastic model
P_{switch}	0.1 W
$P_{\text{conduction}}$	0.06 W
P_{static} (obtained through McPAT)	0.175 W

consumption per clock cycle is characterized at the highest supported V_{DD} level of 1 V. The parameter values of the switching network and the PDN are summarized in Table 5.8.

The additional switching and conduction loss due to the PMOS switches is an insignificant fraction of the total power consumed by the CMP as the losses only occur when the PDN is reconfigured to combine the outputs of the OCVRs. The energy consumption of the proposed power delivery system is up to 44% less than the energy consumed by the CMP with over-provisioned OCVRs. The reduction in energy is primarily due to the optimal PCE offered by the buck converter at the maximum output current supported.

The energy savings with the developed PDN as compared to a conventional PDN with over-provisioned buck converters designed for light load efficiency is also evaluated. Buck converters are designed placing the highest priority to the PCE and improved light load efficiency. The footprint and other figures of merit of the buck converter are, therefore, suboptimal. The properties of the simulated buck converters

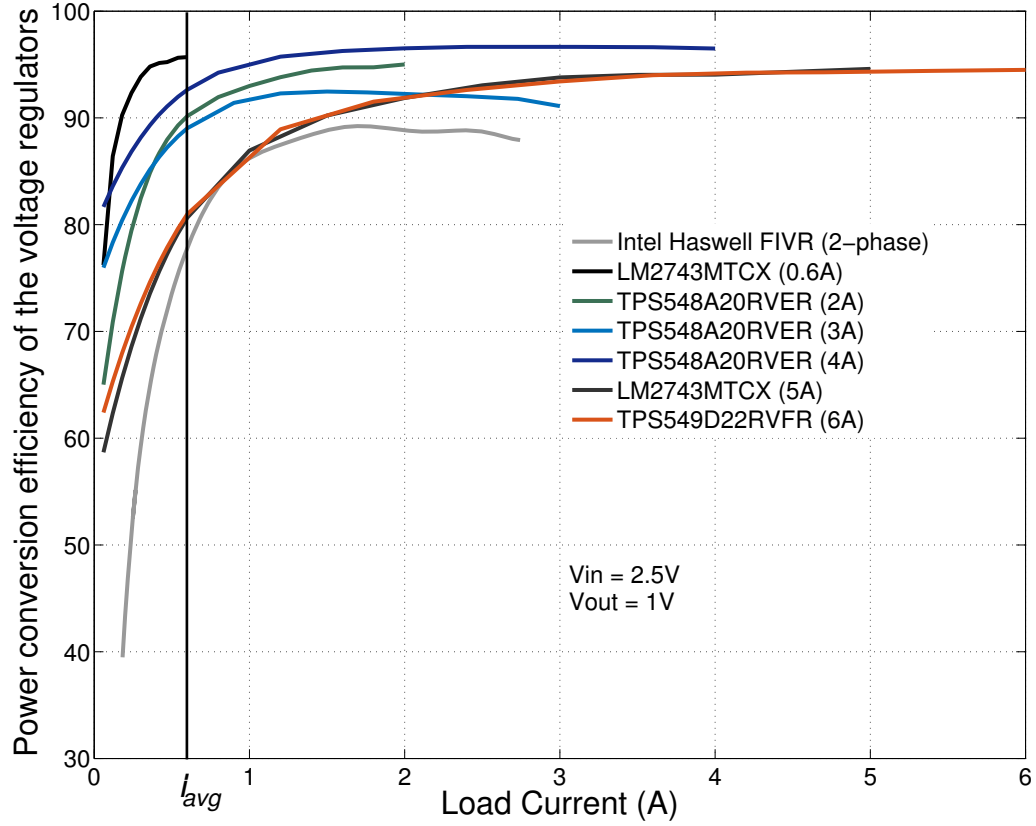


Figure 5.8: Power conversion efficiency of DC-DC switching buck converters simulated with parameters listed in Table 5.7. The model numbers of the devices used for simulation are specified in the figure [106]. The power conversion efficiency of the 2-phase FIVR [31] is provided for comparison with the simulated buck converters that include enhanced light load efficiency.

are summarized in Table 5.7 and the PCE variation with load current is shown in Fig. 5.8.

The energy consumption is characterized for 8, 16, 32, 64, 128, and 256 core CMP systems and OCVR peak current ratings of 0.6 A, 2 A, 3 A, 4 A, 5 A, and 6 A. The core configuration is listed in Table 5.1. The static power losses P_{static} , switching power losses P_{switch} , and conduction power losses $P_{conduction}$ are listed in Table 5.8. The dynamic power consumption $P_{dynamic}$ per core is statistically modeled as a Type

IV Pearson distribution given by (5.1) [105]. The energy consumption of the CMP is computed with (5.4) for a given number of cores and OCVR peak current rating. The percentage savings in energy consumption in comparison to a conventionally over-provisioned CMP for different core counts and OCVR ratings is shown in Fig. 5.9. The larger energy consumption of the CMP system with a peak OCVR rating of 5 A as compared to the 6 A rated OCVR is due to higher reconfiguration of the RPDN, which leads to increased switching and conduction losses in the HSS network. The energy savings for a given OCVR rating do not change significantly with the scaling of the number of cores in the CMP since the dynamic power consumption of the cores is stochastic and the ratio of OCVRs clustered or declustered at run time with respect to the total number of OCVRs in the system remains constant. The ratio of the number of cores consuming current below I_{avg} , above I_{avg} , and in idle state (the j:k:l ratio) does not change significantly.

Simulations with a stochastic model of the dynamic current indicate that the proposed RPDN offers an average reduction in the energy consumption of 15% as compared to a CMP with over-provisioned OCVRs designed with enhanced light load efficiency. With DVFS, the reduction in consumed energy increases proportionally with the scaled voltage and frequency.

5.6 Transient analysis of the RPDN

A reconfigurable power delivery network for integrated circuits has been proposed in [108] using micro-electro-mechanical (MEM) switches. The topology proposed in [108] is suitable for power gating certain sections of the PDN but does not provide

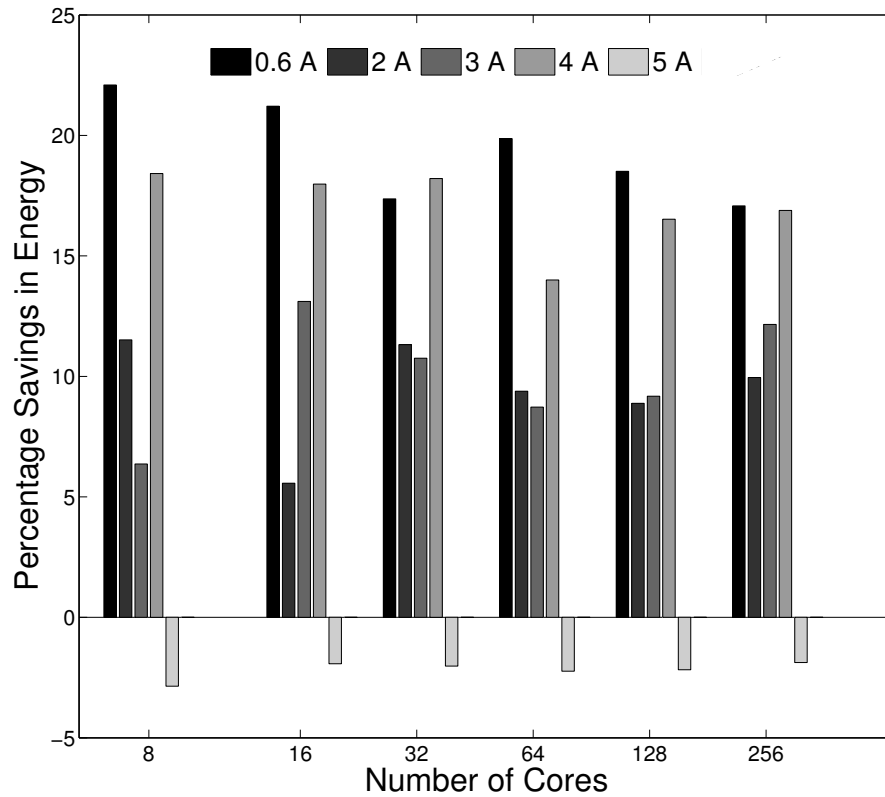


Figure 5.9: Percentage energy saved with the proposed RPDN as compared to a baseline OCVR configuration of 6 A with enhanced light load efficiency. The energy savings is determined for the number of cores in the CMP and for the varying voltage regulator peak current rating of the DC-DC switching buck converters described in Table 5.7.

a fast reconfiguration of the PDN as the switching speed of the MEM switches is three orders of magnitude slower than high performance CMOS switches. Nano-electro-mechanical switches (NEMS) currently offer a switching speed in the range of 10 to 20 ns [109]. However, for the high speed switching network proposed for the load balanced reconfigurable PDN, NEM switches are still an order of magnitude slower than MOS switches. MOS switches, however, introduce finite on-state resistance and off-state leakage. As shown in Section 5.5, despite the static and dynamic power

Table 5.9: IBM power grid benchmark for transient analysis (*ibmpg1t*) [110]. The benchmark includes the number of current sources I , nodes N , resistors R , inductors L , capacitors C , shorts S (zero value resistors and voltage sources), and connections V to the 1.8 V off-chip voltage source. The V off-chip connections are replaced by a Verilog-A model of the OCVR with V connections to the on-chip PDN.

#I	#N	#R	#L	#C	#S	#V
10774	39681	40801	277	10774	14208	100

loss of the MOS switches, the energy efficiency of the multi-core system improves by implementing the developed reconfigurable PDN with under-provisioned OCVRs. In this section, the impact of the MOS switches on the power supply voltage droop when the outputs of two or more OCVRs are combined to provide higher than average current is analyzed.

The PDN selected for characterization of the power supply voltage droop is the two metal layer *ibmpg1t* time domain analysis benchmark circuit released by IBM [111] as part of the 2012 TAU Power Grid Simulation Contest [110]. The configuration of the PDN is summarized in Table 5.9. The power grid metal lines are modeled as resistive elements. The C4 bumps are modeled as inductances of 1 nH in series with a resistance of 0.25 Ω . A spatially distributed current load is modeled with SPICE current pulse sources. Decoupling capacitors along with the effective series resistance of the capacitors are modeled proportional to the pulsed current load [111]. The typical period of the pulsed current sources is 3 ns, and the peak current drawn is 25 A. The passive elements (R , L , and C) and the structure of the *ibmpg1t* benchmark PDN is not modified for the reconfigurable PDN with run-time OCVR clustering.

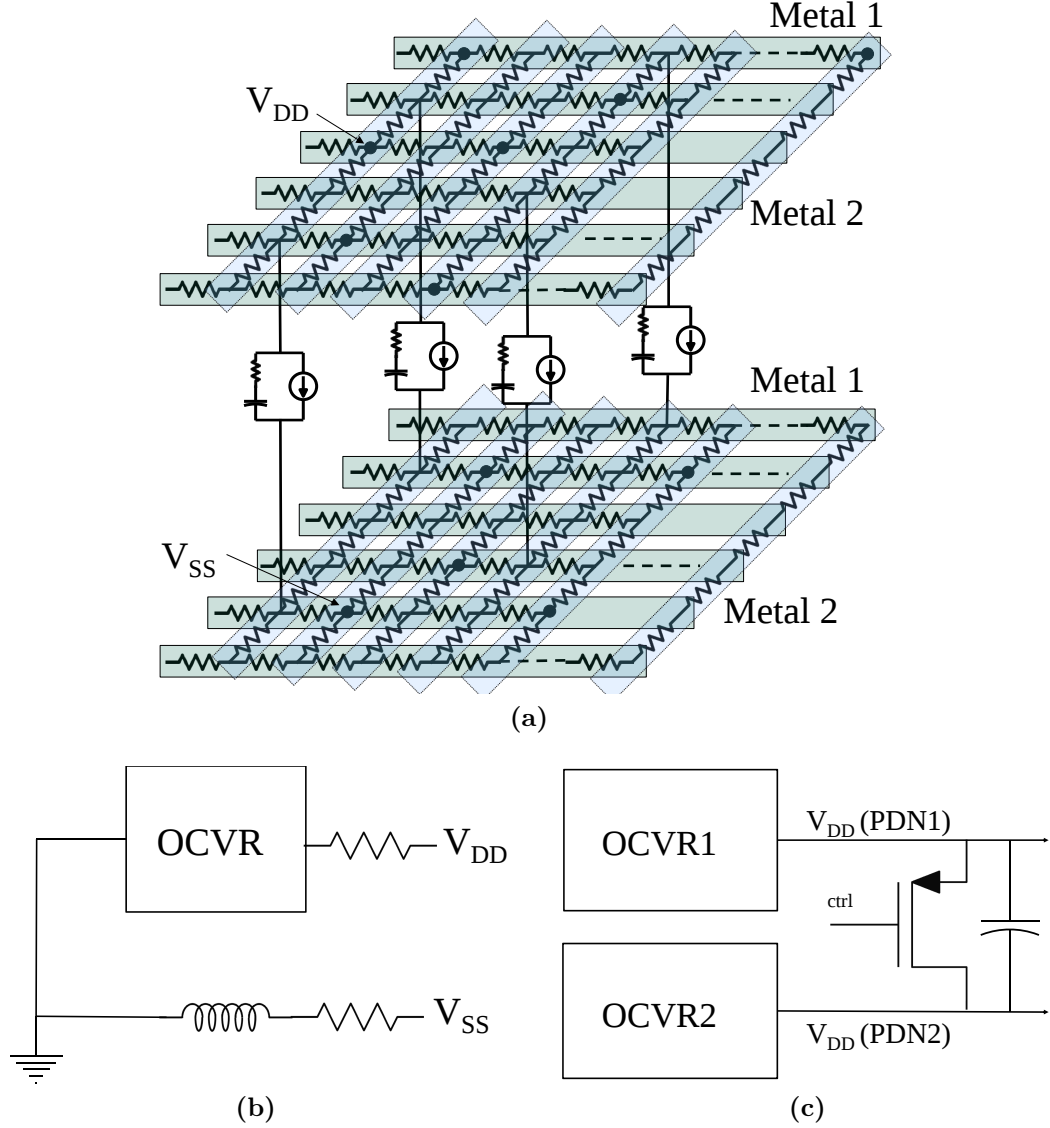


Figure 5.10: Structure of the reconfigurable PDN for transient analysis including (a) a schematic of a section of the PDN based on the IBM Power Grid Benchmark *ibmpg1t* with interdigitated V_{DD} and V_{SS} nets shown on different planes for the same metal layers, (b) a schematic representation of the Verilog-A model of the current limited OCVR connected to the V_{DD} nets of the modified *ibmpg1t* PDN, and (c) the outputs of two Verilog-A models of the current limited OCVR connected through a PMOS switch in an IBM 180 nm technology.

A Verilog-A model of the current limited voltage regulator is developed with a maximum current rating of 15 A. The voltage regulator model includes load current sense and voltage droop control, which are features of commercial buck converters

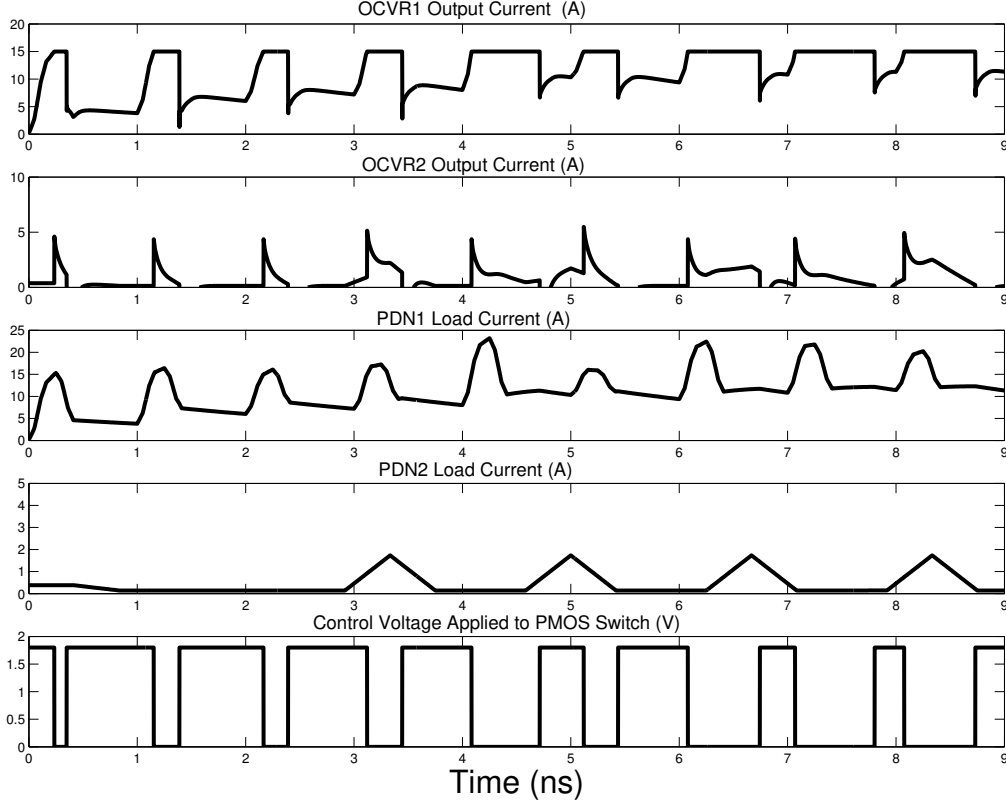


Figure 5.11: Transient analysis of two power distribution networks (modified *ibmpg1t* power grid benchmark) with a peak load current consumption of 25 A and 2 A served by a cluster of two OCVRs with a maximum load current rating of 15 A.

[112]. The output node of the Verilog-A model is connected to 100 voltage source nets across the *ibmpg1t* grid. The connections to C4 bumps are removed as the Verilog-A model emulates an on-chip voltage regulator. The schematic representation of the SPICE model of the *ibmpg1t* PDN and the Verilog-A model of the OCVR connected to the PDN are shown in Fig 5.10.

Two PDNs that include OCVR models are connected through a PMOS switch using an IBM 180 nm technology, as shown in Fig. 13(c). The 180 nm technology node was chosen to match the 1.8 V power supply voltage (V_{DD}) of the *ibmpg1t* benchmark

PDN. A decoupling capacitor is added in parallel to the PMOS switch to reduce the impact of the current transient generated when the switch is turned on. The total decoupling capacitance for the original *ibmpg1t* benchmark PDN is 1.63 μF while the developed PDN with PMOS switches includes 1.79 μF of decoupling capacitance, an approximately 10% increase. Instead of a single PMOS device per OCVR, multiple PMOS switches are connected to voltage source nets on the *ibmpg1t* grid along with a corresponding decoupling capacitor to suppress the inrush current. The peak load current demand on PDN2 is set to 2 A to provide time intervals when the output currents of the two OCVRs are combined through the PMOS switch, which occurs when OCVR1 is not able to supply the total current demand on PDN1.

The current demand of the two PDNs and the current supplied by the respective OCVRs connected to the PDNs is shown in Fig. 5.11 for a time interval of 9 ns. The voltage signal controlling the gate of the PMOS switch connecting the output of the two OCVRs is also shown in Fig. 5.11. When the current consumption of PDN1 exceeds 13 A (the value of ΔI using *Algorithm 5.3a*), the PMOS switch is turned ON, and the output current of the two OCVRs is combined. The combined outputs of the two OCVRs provide uninterrupted current flow to the two PDNs. The impact on the power supply voltage droop on PDN1 while combining the outputs of the OCVRs is analyzed. The time stamp corresponding to the maximum IR drop on any node in PDN1 is identified and the power supply voltage profile across all nodes for that time instant is compared with an unmodified *ibmpg1t* power grid with off-chip voltage regulation. The comparison of power supply voltage noise for both the proposed and unmodified PDNs is shown in Fig. 5.12. Despite the activation and deactivation

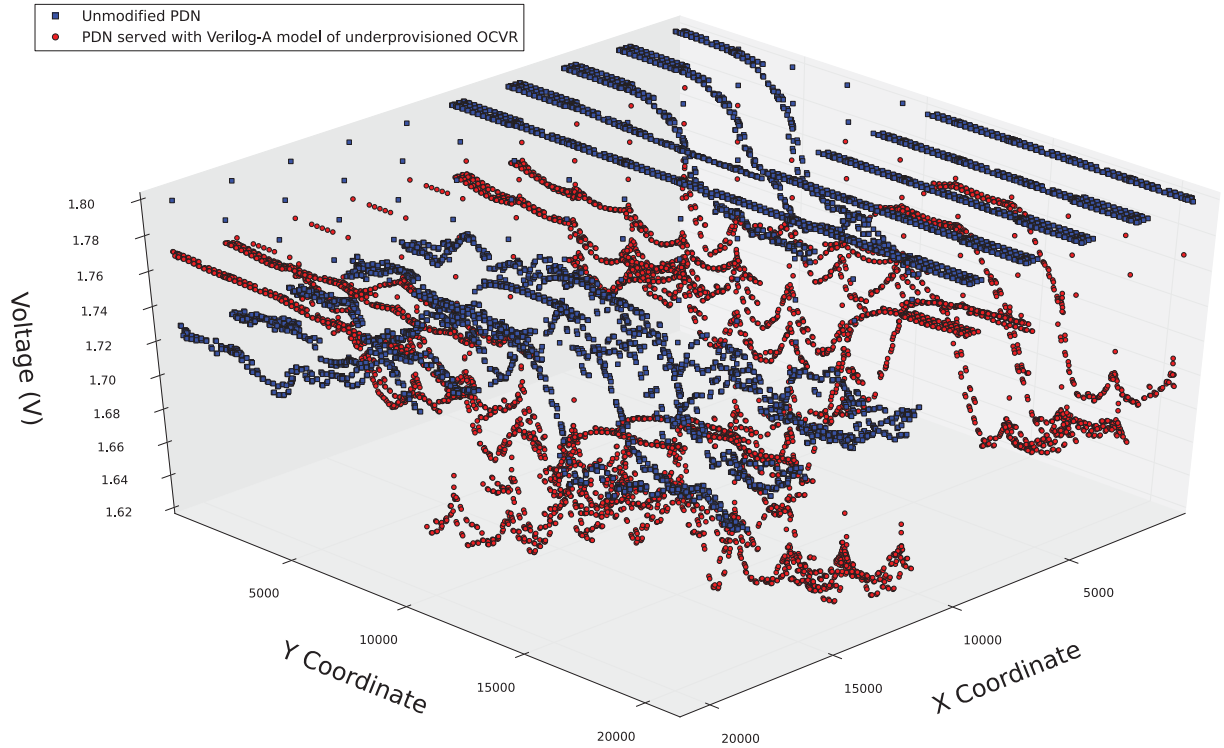


Figure 5.12: Comparison of the worst case voltage droop across the PDN for an unmodified *ibmpg1t* PDN benchmark and a PDN served with the interconnected current limited OCVRs combined at run time.

of the PMOS switches, the maximum IR drop across PDN1 does not exceed 10% of the power supply voltage (1.8 V) and is 0.18% greater than the worst case IR drop measured on the unmodified *ibmpg1t* benchmark PDN with off-chip voltage regulation. A transient analysis of the *ibmpg1t* PDN with a Verilog-A model of the OCVR is performed to demonstrate the feasibility of the proposed reconfigurable PDN for a CMP or many-core system. The analysis is performed on two PDNs, each supplying current to a core. Therefore, the maximum current capability of the OCVR, the width of the PMOS switch, and the additional decoupling capacitance are much larger than that needed when the reconfigurable PDN with OCVRs is designed

for a system with larger core count.

5.7 Summary

A load balanced circuit technique to deliver average power through on-chip voltage regulators (OCVRs) is developed. The current rating of each OCVR is reduced to support only the average current demands of typical workloads executed on the CMP system. The reduction in the maximum output current of the OCVRs improves the power conversion efficiency of the OCVRs, reduces the footprint of the PDN by at least 23%, and improves the energy efficiency of the CMP system by up to 44%. The developed interconnected power delivery network (PDN) is applicable to any OCVR circuit topology and offers higher reliability through a run-time OCVR clustering technique that prevents system failure when the current demand of the core exceeds the maximum output current supported by the OCVRs. The simulated results indicate that the optimum OCVR configuration for a CMP system depends on the average load current requirement per core. The inter-connected power delivery system is applicable to any OCVR circuit topology and offers higher reliability through a run-time clustering technique that prevents system failure when the current demand of the core exceeds the maximum output current supported by a single OCVR.

Chapter 6: Work load scheduling for chip multiprocessors with under-provisioned power delivery

“Sooner or later, the worst possible set of circumstances is bound to occur.”

– Sodd’s second law.

“Any system must be designed to withstand the worst possible set of circumstances.”

– Sodd’s second law, Corollary 1.

Energy efficiency has emerged as a critical design parameter in multi-core or chip multi-processor (CMP) systems. Apart from increased energy efficiency, power delivery through on-chip voltage regulators (OCVRs) offers several benefits including reduced latency to apply DVFS, point of load power delivery with minimal power supply noise, and reduced I/O pin count devoted to power and ground signals [18, 95]. The peak power consumption and worst case power supply noise transient in a CMP determine the design of the power delivery system. Conventionally, the power rating and the design topology of the OCVR is selected based on the maximum possible power consumption of the load circuit. The work done in [15, 17] demonstrates that the conventional method to design the power delivery network results in an OCVR with a maximum output current rating that is over-provisioned by at least an order of magnitude. By under-provisioning the OCVRs to meet the typical or average current load of the circuit, the energy efficiency of the CMP system is increased by up to 44% [17]. A reconfigurable power delivery network (RPDN) with run-time clustering of the outputs of the OCVRs is described in chapter 5. The connections between the

voltage regulators and the cores is managed as described by algorithm 5.3a in chapter 5. The run-time clustering of the voltage regulators is an example of supply side load management. The on-chip power management unit reconfigures the connections between the OCVRs and the cores to meet the changing load current demands of the cores. The run-time reconfiguration of the power delivery network operates under a power constraint: The total power demanded by the cores at any time instant is less than the total power supplied by the OCVRs. The power constraint is expressed mathematically by (6.1) for a CMP with N cores and N OCVRs. I_{sense_x} and V_x are, respectively, the sensed load current and the operating voltage of each core x . I_{avg} and V_{dd_m} are, respectively, the maximum output current and the maximum supported power supply voltage level of each OCVR.

$$\sum_{i=1}^N V_i \cdot I_{sense_i} < N \cdot V_{dd_m} \cdot I_{avg} \quad (6.1)$$

In this chapter, an energy optimized work load scheduling technique is developed that relaxes the power constraint given by (6.1) on the run-time OCVR clustering algorithm developed in chapter 5 [113]. Low power workload scheduling on heterogeneous processors is a widely researched field [89, 114], although the workload schedulers described in literature are oblivious to the power loss of the DC-DC converters delivering regulated power to the cores. The workload scheduler described in this chapter is a demand side load management technique. Workload schedulers are classified into three categories [115]: best effort scheduling, with acceptance test, and robust scheduling. The proposed heuristic imposes an acceptance test on each incom-

ing task in the system and schedules the task on to one of the cores only if the power constraint of the under-provisioned power delivery system is met. Workloads running on a CMP system are either controllable loads with soft deadlines or non-controllable loads with hard deadlines. The rescheduling of controllable tasks reduces the energy consumption of the CMP for a given scheduling cycle. Real time applications fall under the category of non-controllable loads as hard or firm deadline is imposed. In the case of non-real time tasks with soft deadlines and fixed priority, the tasks that violate the power constraint given by (6.1) are executed in the next scheduling cycle, leading to a performance penalty.

The energy consumed by a taskset on a processing element is a convex function of the computational capacity of the processing element and the task execution time. A convex energy optimization problem is solved to ensure the reliability of the proposed reconfigurable power delivery system with under-provisioned on-chip voltage regulators. The optimization problem is constrained by the total power budget of the CMP and is limited to the peak current rating of the OCVRs. The feasibility of the solution, determined by solving the optimization problem, is demonstrated through a real time workload scheduling heuristic. The scheduler is applicable to homogeneous and heterogeneous CMPs.

The rest of the chapter is organized as follows: The models developed for the CMP system, OCVRs, real time periodic taskset, and power consumption of the cores are described in Section 6.1. The convex energy optimization problem for workload scheduling is discussed in Section 6.2. An energy efficient workload scheduling heuristic is described in Section 6.3. The evaluation of the workload scheduling heuristic

Table 6.1: Frequency (MHz) and voltage (V) pairs used by the DFVS procedure in *Algorithm 6.3a*. The bold values listed in the table are the nominal frequencies and voltages.

big core (A15) (MHz/V)	LITTLE core (A7) (MHz/V)
1800/1.250	1200/1.225
1700/1.200	1100/1.125
1600/1.162	1000/1.100
1500/1.137	900/1.037
1400/1.100	800/0.987
1300/1.062	700/0.950
1200/1.025	600/0.950
1100/1.000	500/0.950
1000/0.962	400/0.950
900/0.925	300/0.950
800/0.900	200/0.950

tic on homogeneous and heterogeneous CMP platforms is provided in Section 6.4. A summary of the work is provided in Section 6.5.

6.1 System model and notations

The under-provisioned CMP includes a set of processing elements or cores and per-core on-chip voltage regulators. The models constructed for the core architecture, CMP platform, voltage regulators, real time periodic tasks, and the power consumption of the cores are described in Subsections 6.1.1 through 6.1.4.

6.1.1 CMP models

CMP systems with homogeneous and heterogeneous core are developed to analyze the work load scheduler. The homogeneous CMP includes processing elements based on the ARM A15 core integrated in the Samsung Exynos 5410 platform [116, 117].

Table 6.2: Parameters of the CMP cores derived from the Samsung Exynos 5410 big.LITTLE architecture [116].

Parameter	big core (A15)	LITTLE core (A7)
Nominal voltage (V)	1.16	1.225
Nominal frequency (MHz)	1600	1200
OCVR maximum current rating (I_{avg} in mA)	800	110
Power model parameters		
α_j	2.63	3.28
κ_j (mW/MHz ³)	2.91×10^{-6}	1.00×10^{-8}
β_j (mW)]	146.49	34.24
Number of cores in homogeneous CMP system	16	0
Number of cores in heterogeneous CMP system	4	4

The parameters used in constructing a 16 core homogeneous CMP platform are listed in Table 6.2. An eight core heterogeneous CMP with four ARM A15 and four A7 cores from the Exynos 5410 platform is also evaluated. The frequencies and voltage pairs applied to the cores for DVFS are listed in Table 6.1. The variation in the power consumption of the core with frequency, based on the power model given by (6.2) and validated in [118], is shown in Fig. 6.1.

6.1.2 Power model

The power consumption of a processing element π_j is approximated as a function of frequency, similar to work done in [118]. The power consumed by any processing element is given by (6.2). The $\kappa * f^\alpha$ and β terms in (6.2) represent, respectively, the dynamic and static power consumption of the cores. The model parameters κ , α , and β for the Samsung Exynos A15 and A7 processors [118] are used to validate *Algorithm 6.3a*. The power consumption with frequency using the estimated model

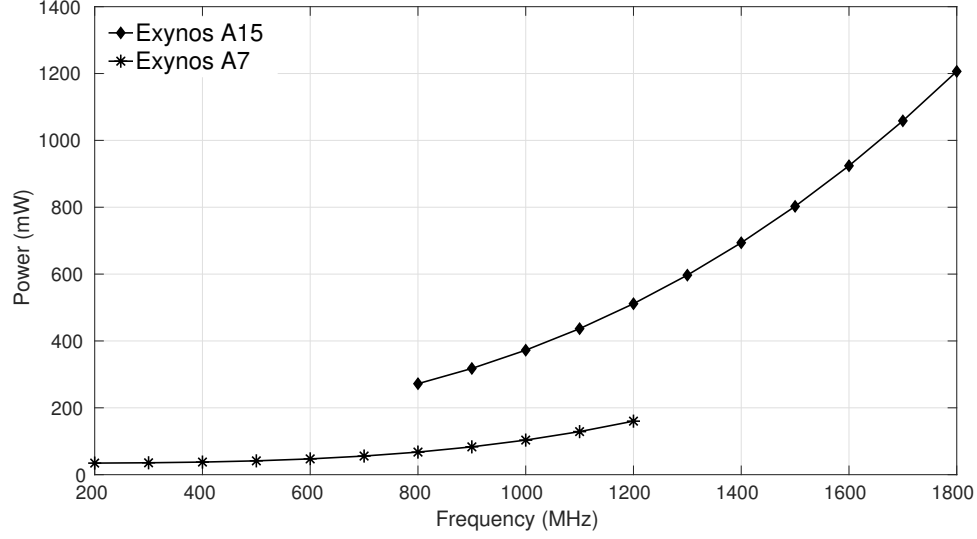


Figure 6.1: Power consumption of the Exynos big.LITTLE cores with frequency based on the model given by (6.2). The power model parameters are validated in [118].

parameters is shown in Fig. 6.1.

$$P(f) = \kappa * f^\alpha + \beta \quad (6.2)$$

6.1.3 Voltage regulator models

An on-chip power delivery network with per core voltage regulation is considered for the CMP system. The on-chip voltage regulators (OCVRs) are modeled as DC-DC switching buck converters [24]. Buck converters with optimum power conversion efficiency based on the power consumption of the A15 and A7 cores, as shown in Fig. 6.1, are developed using [106]. The parameters of the buck converters considered for both the big and LITTLE cores are listed in Table 6.3. The variation in power conversion efficiency with load current is shown in Fig. 6.2.

Table 6.3: Operating values of the switching DC-DC buck converters [106] serving the ARM A15 and A7 cores.

Device	LM3671TLX	TPS62097RWKR
Input voltage (V)	2.5	2.5
Output voltage range (V)	0.9 to 1.3	0.9 to 1.3
Maximum output current (mA)	110	800
Efficiency at maximum output current	90.7	88.9
Peak to peak inductor ripple current (mA)	146.93	357.24
Switching frequency (MHz)	2	1.69
Duty cycle (%)	46.18	47.44
Peak-to-peak output ripple voltage (mV)	1.322	3.215
Total power dissipation (mW)	12.4	109.88
Footprint (mm^2)	37	93

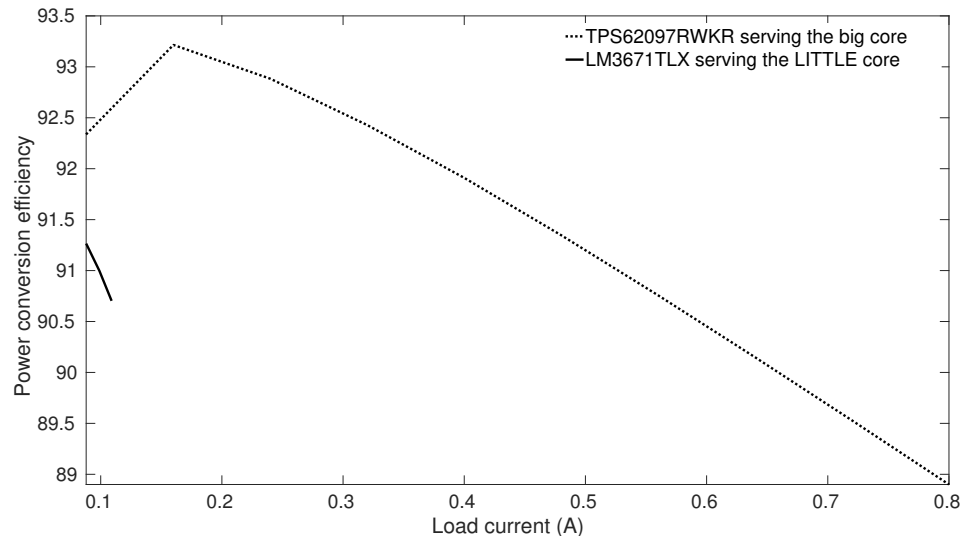


Figure 6.2: Power conversion efficiency of the voltage regulators serving the Exynos A15 (big) and Exynos A7 (LITTLE) cores [116].

6.1.4 Real-time periodic task model

The real-time workloads are modeled as a set of independent periodic tasks $\tau_i \in \mathcal{T}$ to be scheduled on a subset of cores of a many-core system $\pi_j \in \Pi$ [118]. Each task τ_i , has a hard deadline of D_i . Each core π_j supports distinct DVFS levels with voltage $V_x \in [V_{dd_1}, V_{dd_2}, \dots, V_{dd_m}]$ and frequency $f_x \in [f_1, f_2, \dots, f_m]$. A task τ_i with a hard deadline D_i requires at most $C_{i,j}$ cycles to execute on a core π_j at the highest supported voltage V_{dd_m} and frequency f_m . The context switching overhead and the overhead due to resource sharing amongst tasks that remain unresolved after task partitioning are included in $C_{i,j}$. The computational capacity required by task τ_i on core π_j is defined as $u_{i,j} = \frac{C_{i,j}}{D_i}$. The subset of tasks T_j that are executed on core π_j therefore require a total computational capacity of $U_j = \sum_{\tau_i \in T_j} u_{i,j}$ cycles per second.

6.2 Optimal workload scheduling

An optimization problem is defined to partition and schedule real time workloads on a many-core platform. A specific set of constraints unique to the proposed reconfigurable PDN are considered, which account for the use of under-provisioned on-chip voltage regulators. The objective of the optimization problem is to minimize the energy consumption of the many core platform, including the power consumed by the OCVRs. The energy consumed by the system in a given scheduling period T_{epoch} is given by (6.3), where $P(U_j)$ is the power consumed by the core π_j with computational capacity U_j to execute the scheduled task set, and PCE_{U_j} is the combined power conversion efficiency of the OCVR(s) supplying current to the core π_j . The

workload scheduling is constrained by the total computational capacity available to execute the taskset U_j on π_j , where the total capacity must exceed the computational demand of the taskset as given by (6.4). In addition, the operating frequency must fall within the supported frequency range of the cores as given by (6.5). The total power consumed by the cores at any time instant must be less than the combined maximum power supported by all OCVRs in the system as described by (6.6).

$$\min_{U_j} \sum_{\pi_j \in \Pi} \frac{P(U_j)}{PCE_{U_j}} \cdot T_{epoch} \quad (6.3)$$

$$\text{s.t.} \quad \sum_{\pi_j \in \Pi} U_j \geq \sum_{\tau_i \in T} u_i \quad (6.4)$$

$$f_{1,j} \leq U_j \leq f_{m,j} \quad \forall \pi_j \in \Pi \quad (6.5)$$

$$\sum_{\pi_j \in \Pi} P(U_j) < N \cdot V_{dd_m} \cdot I_{avg} \quad (6.6)$$

6.3 Workload scheduling heuristic

A heuristic is described in this section, that performs the real time workload scheduling on the cores for the optimization problem developed in Section 6.2. The heuristic consists of three procedures: *PARTITION*, *DVFS*, and *SCHEDULE*. The *PARTITION* procedure is an evolution of the *Marginal-Power Heuristic (M-PWR)* developed in [118]. An optimal workload partitioning is achieved by incrementing the load on each core such that the constraint given by (6.5) is not violated. The tasks $\tau_i \in \mathcal{T}$ are first sorted in decreasing order of the maximum computational demand

$u_{i,j}$ on cores $\pi_j \in \Pi$. A task is assigned to a core if the scheduling of the task results in the least increase in the power consumption. The output from the procedure is a scheduled taskset Θ_j on each core.

Algorithm 6.3a Real time workload partitioning and scheduling on a many-core system with under-provisioned OCVRs.

Inputs:

Set of real time tasks: $\mathcal{T}$

Set of N cores in the many-core system: Π

Outputs: Schedulable taskset (Θ_j) on each core $\pi_j \in \Pi$ with assigned voltage $V_j \in [V_{dd_1}, V_{dd_2}, \dots, V_{dd_m}]$ and frequency $f_j \in [f_1, f_2, \dots, f_m]$

procedure PARTITION($\mathcal{T}, \Pi$)

for each $\pi_j \in \Pi$ **do**

$\Theta_j \leftarrow \emptyset, U \leftarrow 0$

$\mathcal{T}' \leftarrow \text{SORT}(\mathcal{T} \text{ by descending } \max_j u_{i,j})$

for each $\tau_i \in \mathcal{T}'$ **do**

$\Pi' \leftarrow j: U_j + u_{i,j} < f_{m,j} \triangleright$ Cores on which τ_i is schedulable

if $\Pi' = \emptyset$ **then return** Failed to schedule

$k \leftarrow \arg \min_{j \in \Pi'} P_j(U_j + u_{i,j}) \triangleright$ core id on which τ_i consumes least power

$\Theta_k \leftarrow \Theta_k \cup \tau_i \triangleright \Theta_k$ is the schedulable set of tasks on π_k

$U_k \leftarrow U_k + u_{i,j}$

return Θ

procedure DVFS(Θ)

while $\sum_{\pi_j \in \Pi} f_j^\alpha > (P_{total} - N \cdot \beta) / \kappa$ **do**

for each $\pi_j \in \Pi$ **do**

while $\sum_{\tau_i \in \Theta_j} u_{i,f_j} \leq f_m$ **do**

$f_j \leftarrow (f_x, f_x \in F \text{ and } f_x \leq f_j) \triangleright$ lower the operating frequency to one of the supported DVFS levels $F = (f_1, f_2, \dots, f_m)$

procedure SCHEDULE(Θ_j, f_j)

$k \leftarrow \arg \min_{\tau_i \in \Theta_j} D_i$

$\pi_j \leftarrow \tau_k$

The *DVFS* procedure reduces the operating frequency and the voltage of the cores until the constraint given by (6.6) is satisfied. The right hand side of (6.6) is a constant value equal to the total power P_{total} of the CMP. Expressing the total power consumed by the cores with the power model given by (6.2) in (6.6) constrains the operating frequency of the cores raised to the power α as given by (6.8). The use of

Table 6.4: Parameters to generate real time periodic tasksets.

Parameter	Value
Number of tasks (N_t)	[32, 48, 64, 80]
Task utilization range	[0.1 to 0.9]
Task period range in seconds (T_i)	[10 to 100]
Taskset utilization factor (ρ)	[0.1 to 1]

the *DVFS* procedure results in the optimal frequency of operation for each core by solving the bounded knapsack problem. The deadline of each task in the taskset Θ_j is analogous to the value of the item in the knapsack. The required computational demand at a given frequency f_j on processor π_j is u_{i,f_j} . The weight added to the knapsack is analogous to u_{i,f_j} . The objective of the knapsack problem is to maximize the number of tasks executed on a core without violating the task deadline. The *DVFS* procedure lowers the operating frequency of each task until constraints (6.7) and (6.8) are satisfied. Once the operating frequency of each task in Θ_j is determined, the *SCHEDULE* procedure schedules the tasksets on each core based on an earliest deadline first policy.

$$\sum_{\tau_i \in \Theta_j} u_{i,f_j} \leq f_m \quad (6.7)$$

$$\sum_{\pi_j \in \Pi} f_j^\alpha \leq (P_{total} - N \cdot \beta)/k \quad (6.8)$$

6.4 Simulation results

Real time periodic tasks with implicit deadlines are considered to characterize the efficacy of the proposed task scheduler. The task scheduling is performed for one

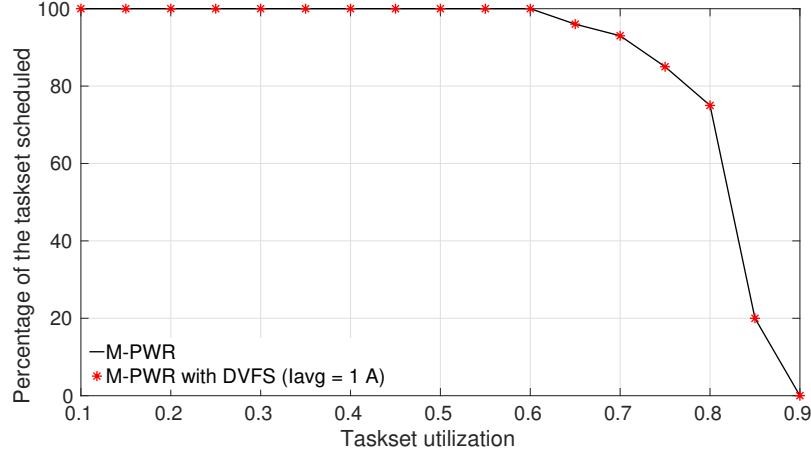


Figure 6.3: Percentage of tasks successfully partitioned by the M-PWR heuristic [118] and successfully scheduled by *Algorithm 6.3a* in a 16 core homogeneous CMP. Task scheduling through *Algorithm 6.3a* matches the M-PWR heuristic with the maximum output current rating of the voltage regulator set to 1 A.

hyper period of the taskset T_{epoch} , which is the least common multiple of the implicit deadlines of all tasks $\tau_i \in \mathcal{T}$. The tasks are generated with the parameters listed in Table 6.4. The computational capacity u_{ij} of the tasks is selected as a random variable with a uniform distribution between 0.1x to 0.9x the maximum supported operating frequency of the cores in the CMP (maximum frequency f_m of 1800 MHz). The total computational time requested by the taskset in a hyper period is less than the available time on the processing elements to prevent system overload, ensuring that the taskset utilization factor or the system load is less than 1 ($\rho < 1$).

The resulting task schedule, from execution of *Algorithm 6.3a* on a homogeneous CMP platform with 16 cores configured as Exynos 5410 A15s, is shown in Fig. 6.3. The task scheduling is constrained due to the limited power budget of the under-provisioned voltage regulators. For a maximum output current I_{avg} of 1 A, the percentage of tasks scheduled by *Algorithm 6.3a* is identical to the M-PWR heuristic

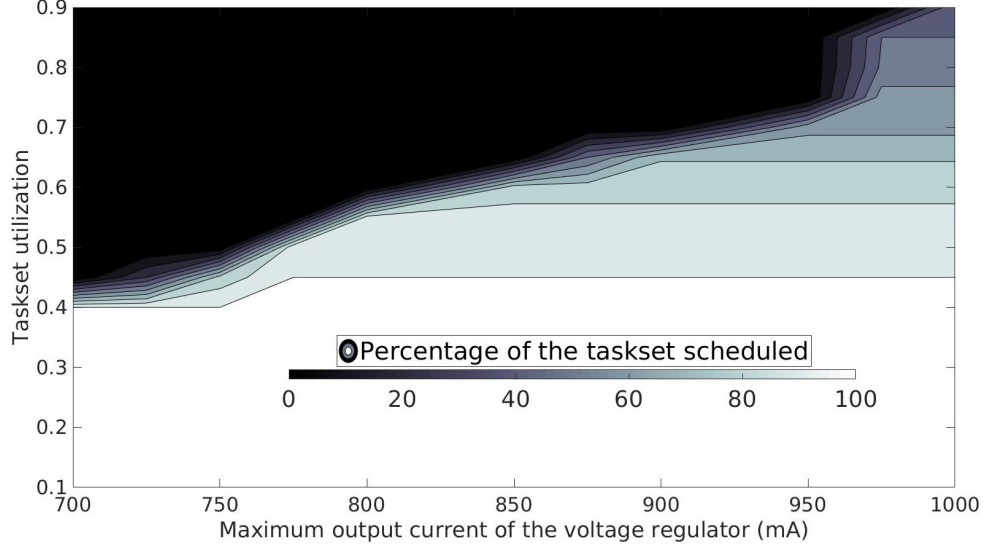


Figure 6.4: Contour plot of the percentage of tasks successfully scheduled by *Algorithm 6.3a* with varying taskset utilization and maximum output current of the voltage regulators in a homogeneous CMP with 16 cores.

in [118]. The execution of *Algorithm 6.3a* is further characterized on a homogeneous platform with voltage regulators of varying maximum output current I_{avg} . The results are shown through the contour plot in Fig. 6.4. The percentage of tasks scheduled for a given taskset utilization factor decreases as the maximum output current of the voltage regulators is reduced. For a voltage regulator designed with a maximum output current I_{avg} of 0.7 A, the percentage of tasks scheduled matches the M-PWR heuristic [118] up to a taskset utilization factor ρ of 0.65.

The workload scheduler is also evaluated on a heterogeneous CMP platform with four Exynos A15 (big) and four Exynos A7 (LITTLE) cores. For a randomly chosen taskset hyper period, the task distribution and corresponding computational demand ($u_{i,j}$ of each task) is shown in Fig. 6.5. There are 11 tasks assigned to the big core cluster and five to the LITTLE core cluster. The maximum output current of the

voltage regulators serving each of the big cores is set to 800 mA and the voltage regulators serving each of the LITTLE cores to 110 mA. The frequency assigned to each core to meet the constraint given by (6.8) is determined and shown in the Fig. 6.5. Depending on the total computational demand of the tasks assigned to each core, the frequency is lowered from the maximum supported frequency of 1800 MHz for the big cores and 1200 MHz for the LITTLE cores. The task partitioning performed by the *PARTITION* procedure further improves power efficiency by preferentially assigning tasks to the LITTLE cores, that meet the task utilization constraint given by (6.4). Consequently, for an identical scaling factor of the peak output current of the OCVRs (I_{avg}/I_{peak}) serving the LITTLE core and the big core, the percentage of tasks scheduled through the *DVFS* procedure is lower for the LITTLE cores as compared to the big cores. As the LITTLE core cluster has a load current range of 100 mA, the scaling factor of the maximum output current of the voltage regulators serving the LITTLE cores is set to a larger value than that for the big cores to achieve a high task schedulability on the heterogeneous platform.

The task scheduling results for the homogeneous and heterogeneous CMP demonstrate that the proposed workload scheduler in tandem with the runtime on-chip voltage regulator clustering algorithm developed in chapter 5, offer an efficient and robust cross layer energy optimization of CMPs with under-provisioned on-chip voltage regulators.

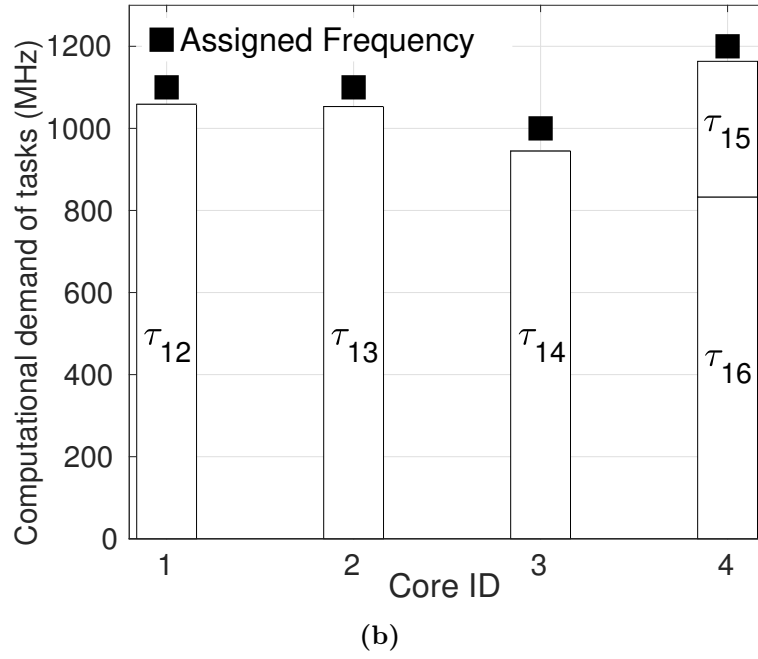
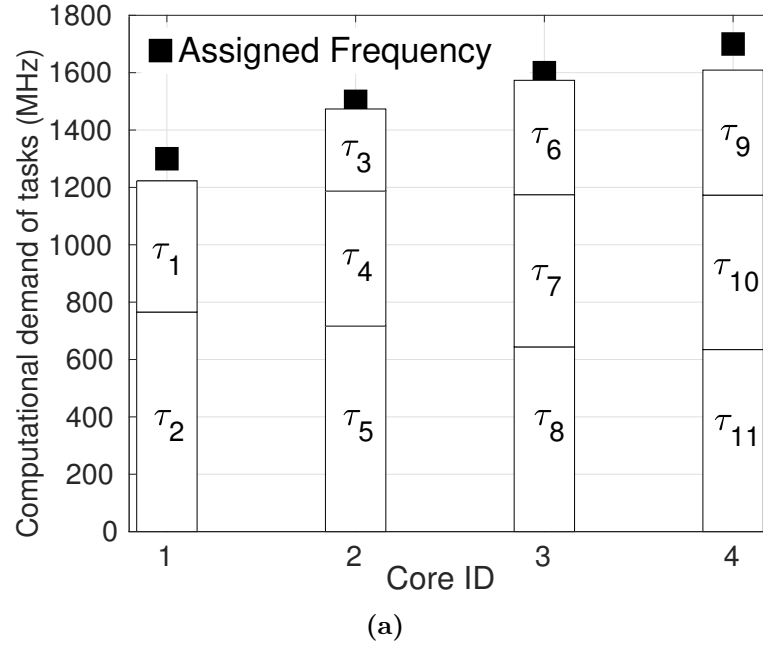


Figure 6.5: Snapshot of the task assignment on a heterogeneous CMP platform with (a) big cores modeled on A15 parameters [116] and (b) LITTLE cores modeled on A7 parameters [116]. The maximum output currents of the voltage regulators serving each of the big cores and LITTLE cores are, respectively, 800 mA and 110 mA.

6.5 Summary

A real-time workload mapping heuristic is developed to minimize the reconfiguration of the power delivery network with under-provisioned on-chip voltage regulators. The scheduled tasks are assigned optimum DVFS levels for each core. The heuristic is evaluated on homogeneous and heterogeneous CMP platforms with real time periodic tasks. The schedulability of the tasks with varying taskset utilization is compared against M-PWR, an energy efficient workload scheduler. There is 100% scheduling of tasksets and assignment of DVFS levels for each core of a homogeneous CMP with task utilization factor of up-to 0.55. The workload mapping heuristic in conjunction with the run-time reconfiguration of the power delivery network ensure reliable and energy efficient operation of the CMP with on-chip voltage regulators designed for only the typical or average load current demand.

Chapter 7: Hyperabrupt varactors for power supply noise mitigation

"It would be a dull gray world without noise."

– Leon Cohen

Robust and efficient on-chip power delivery is a challenge with increasing transistor density due to scaling and limited temperature and power budgets. Power supply noise naturally occurs due to sudden variation in the activity of digital circuits and, therefore the load current drawn from the power distribution network. The on-chip power integrity is maintained through significant voltage margins to negate any timing violations that occur due to large and fast power supply transients. With technology scaling the nominal supply voltage is reduced. However, the threshold voltage has remained relatively unchanged, leading to an increased sensitivity to power supply voltage variation [119]. The power delivery of a Pentium 4 is simulated with predictive technology models (PTM) of transistors from 45 nm to 11 nm. The power supply voltage for each technology node is scaled based on ITRS projections [120]. The on-chip current load was varied from 50 A to 100 A for a 45 nm technology. For each subsequent technology node, the current was scaled inversely to the power supply voltage, resulting in a constant power budget. The worst case peak to peak swing of the power supply voltage with technology scaling is shown in Fig. 7.1 [119]. For the same power budget the power supply noise is projected to double for the 16 nm node as compared to the 45 nm node .

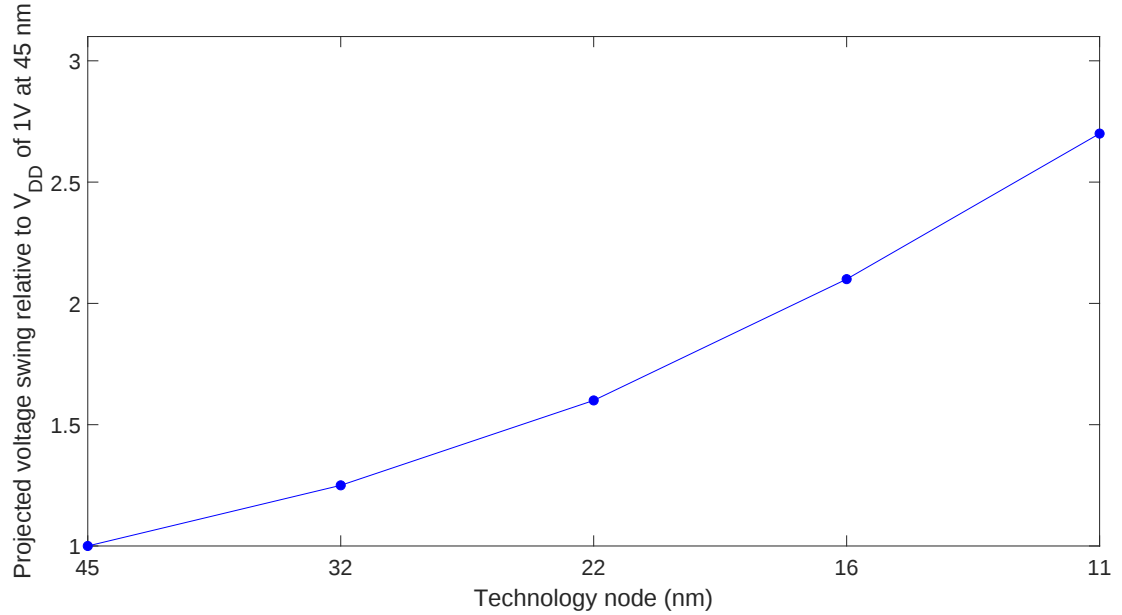


Figure 7.1: Projected power supply noise characterized by voltage swing normalized to a V_{DD} of 1 V for a 45 nm node [119].

The power budgets for multi-core systems implemented in advanced technology nodes increase while only minor improvements in the power delivery network are made due to limitations in material properties. Research to compensate for large transient power supply droops through decoupling capacitor placement and enhancement in chip packaging technology has been completed [121–123]. However, with increasing transistor density and complex workloads executed on a CMP system, power supply droops are inevitable. As the power supply droop impacts the critical path delay, significant voltage guard band is added to the output voltage of the voltage regulators (VRs) to prevent timing and logical failure in the critical path. Power supply droop, therefore has a negative impact on both the performance (limitation in the maximum operating frequency) and the energy consumption (voltage guard band results in higher power consumption). Current methods to compensate for the power

supply droop are reactive in nature. Methods such as critical path monitors [124] and adaptive clocking [125] react to the voltage droop event and are, therefore less effective to offset the impact of a droop.

In this chapter, an on-chip hyperabrupt junction diode, which is a voltage dependent capacitive structure, is explored for power supply noise suppression [126]. The variable capacitance of the device is exploited to suppress power supply undershoots and overshoots. Detailed power supply noise simulations are performed on an optimized PDN to study the on-chip voltage behavior while using hyperabrupt junction varactors as decoupling capacitors. A circuit technique with series connected hyperabrupt junction diodes as decoupling capacitors is proposed to suppress power supply voltage noise.

The remainder of the chapter is organized as follows. The power supply noise induced due to clock-edge triggered digital logic is described in Section 7.1. The physical and electrical properties of a hyperabrupt varactor are described in Section 7.2. SPICE simulations with hyperabrupt varactors as on-chip decoupling capacitors are shown in Section 7.3. The comparison between different power supply noise mitigation schemes with hyperabrupt varactors is done in Section 7.4. The summary of the chapter is provided in Section 7.5.

7.1 Clock-edge induced power supply noise

In a synchronous digital system, the activity of a CMOS circuit is driven by the clock signal. The instantaneous sinking of charge on the clock edge gives rise to an impulse of load current, which results in a voltage change across the inductive path

($L \cdot dI/dt$) of the power distribution network connecting the load circuit to the voltage regulator. Due to the latency of the rate of change of current through the inductance, the current provided through the PDN does not meet the instantaneous impulse load current demand.

Significant work has been completed on decoupling capacitor optimization to limit power supply noise [121,127,128]. Decoupling capacitors act as local charge reservoirs that meet the instantaneous charge requirement of the load circuit. However, the voltage across the decoupling capacitor drops in proportion to the ratio of the charge pulled for the capacitor and the total capacitance, which leads to a voltage droop on the power delivery network (PDN). The impulse of total charge consumed on the clock edge in a CMOS circuit is, therefore, the fundamental cause of the variation in the power supply voltage. The voltage droop is reduced by increasing the value of the decoupling capacitance, but at the cost of increased area and greater leakage current from the charging and discharging of the larger on-chip capacitance.

A switched capacitor load model (refer to Fig. 7.2) is developed to characterize the clock edge induced power supply noise on the PDN and the behavior of the instantaneous current provided by the on-chip decoupling capacitance. The load model emulates CMOS circuit behavior and provides an accurate representation of the voltage variation on the PDN connecting the on-chip capacitance and circuit loads [129]. The variation in load current results in simultaneous changes in PDN voltage, which is not feasible with load current models using ideal constant current sources.

The parameters used to model and simulate the switched capacitor load are listed

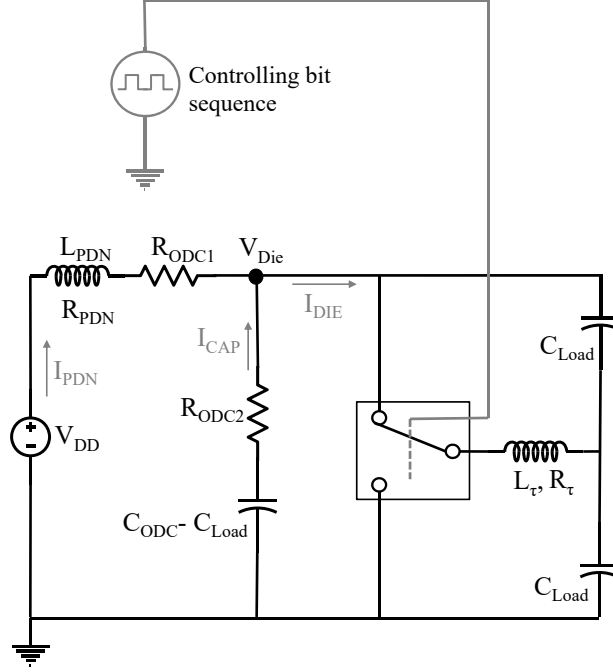


Figure 7.2: Switched capacitor load circuit to emulate the switching behavior of clock edge triggered CMOS circuits. Power supply voltage is set through a PDN with characteristic impedance equal to the target impedance.

in Table 7.1. An ideal 1 V (V_{DD}) voltage regulator provides current through a PDN modeled with a single impedance peak attributed to the package loop inductance, package series resistance, on-chip capacitance, and load circuit. The target impedance of the PDN is chosen to equal the characteristic impedance for a dynamic current load $I_{dynamic}$ of 1.55 A consumed per clock cycle. At a clock frequency f_{clock} of 1 GHz, the charge consumed on each rising and/or falling edge of the clock is 1.55 nC, as given by (7.1). The amount of load capacitance C_{Load} charged and discharged on each clock cycle for a circuit operating at 1 V is therefore 1.55 nF, as given by (7.2). The interconnect inductance L_{τ} along with the corresponding interconnect resistance R_{τ} determine the charge rate of C_{Load} over the entire clock cycle. The interconnect

Table 7.1: Parameters used to simulate the switched capacitor load circuit [129].

Parameter	Value
Power supply voltage, V_{DD}	1 V
PDN characteristic impedance, Z_0	32 m Ω
PDN target impedance, Z_{target}	32 m Ω
PDN loop inductance, L_{PDN}	50.7 pH
PDN loop resistance, R_{PDN}	5.1 m Ω
Clock frequency, f_{clock}	1 GHz
On-chip capacitance, C_{ODC}	50 nF
On-chip resistance due to metallization, R_{ODC1}	3 m Ω
On-chip resistance due to MOSFET transconductance and signal wiring, R_{ODC2}	2 m Ω
Current consumed per clock cycle	1.55 A
Capacitive load switching per clock cycle, C_{Load}	1.55 nF
Inductance controlling charging of C_{Load} , L_τ , R_τ	6.7 pH, 80.65 m Ω

resistance for the given load capacitance is chosen such that the $R_\tau C_{Load}$ time constant τ is one fourth the clock period to ensure that the current impulse is distributed across the entire single clock cycle. The total on-chip capacitance C_{ODC} is 50 nF [129]. Therefore, the intentional on-chip decoupling capacitance is the difference between C_{ODC} and C_{Load} .

$$Q_{clk-edge} = \frac{I_{dynamic}}{f_{clock}} \quad (7.1)$$

$$C_{Load} = \frac{Q_{clk-edge}}{V_{DD}} \quad (7.2)$$

An emulated load current is controlled by an applied bit sequence to a single pole double throw switch, which results in a set charge and discharge time of the switching capacitance C_{Load} . A current impulse is generated by the bit sequence

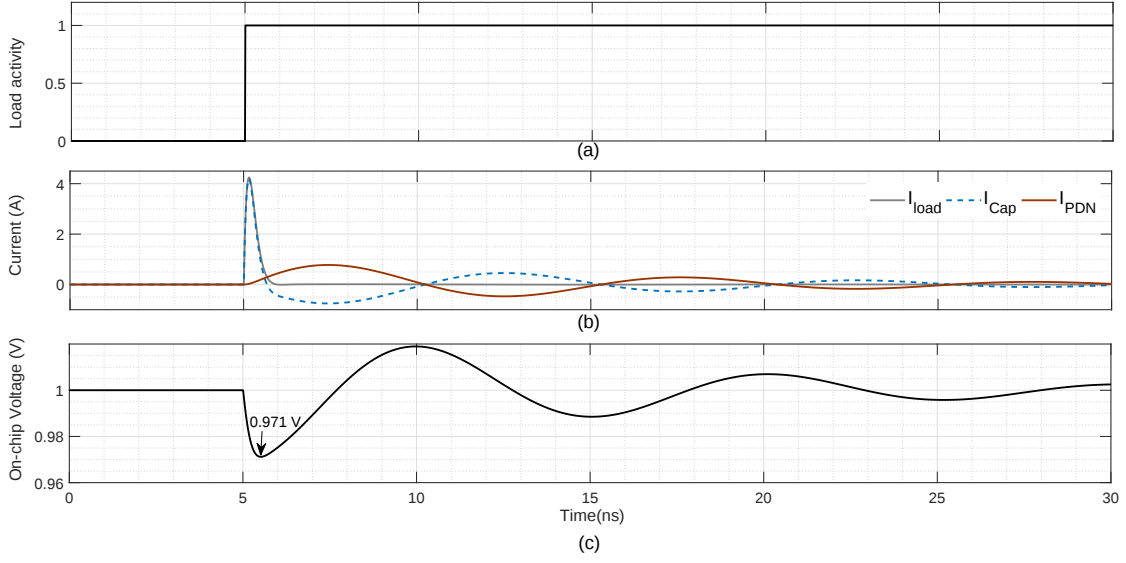


Figure 7.3: An impulse of load current simulated on the switched capacitor load circuit shown in Fig. 7.2. The sub-figures depict (a) the controlling bit sequence implemented as a step function, (b) the resulting impulse of load current (I_{load}), the discharging current from the decoupling capacitor (I_{CAP}), and the current through the voltage source (I_{PDN}), and (c) the on-chip voltage response to the impulse current.

shown in Fig. 7.3(a). The current consumed by the load, the current delivered by the on-chip capacitance, and the current supplied by the voltage source are shown in Fig. 7.3(b). The on-chip voltage variation due to the dynamic current consumption is shown in Fig. 7.3(c). The maximum voltage droop on the on-chip PDN is 29 mV, which is within 7% of the theoretical calculation of 31 mV given by (7.3). The 2 mV difference (reduction) from the theoretical calculation is due to the charge delivered through the PDN inductance.

$$\Delta V_{droop} = \frac{I_{dynamic}}{f_{clock} \cdot C_{odc}} \quad (7.3)$$

7.2 Hyperabrupt junction varactor diode

A reverse biased P-N junction diode acts as a voltage controlled capacitor. The applied reverse bias voltage V_R controls the thickness of the depletion region, which in turn determines the junction capacitance. The capacitance of the junction is inversely proportional to the thickness of the depletion region. P-N junction diodes, which exhibit low losses at microwave frequencies, are manufactured with a controlled doping profile to enhance the variation of the junction capacitance with the applied reverse bias voltage [130]. The diodes, also referred to as *varactors* [130], have been extensively used as radio frequency translation devices. The large tuning ratio of the capacitance, with moderate costs in linearity and quality factor, makes hyperabrupt varactors suitable for many frequency tuning applications [130] as well as more common circuit applications such as voltage controlled oscillators, phase shifters, and frequency multipliers.

A varactor diode generally includes a junction with a heavily doped P side. The variation of the doping concentration on the N side sets the feasible capacitance range of the varactor as a function of the reverse bias voltage. The ideal doping profiles of three P+/N junctions are shown in Fig. 7.4a. The doping concentration on the P+ side is a constant ($N = N_a$). The doping concentration on the N side is mathematically expressed as $N_d \propto x^{-m}$ for $x > 0$, where x is the distance from the junction and m is a constant that describes the type of P-N junction. For an abrupt varactor and linearly graded varactor, m is 0 and 1, respectively. The m value for a hyperabrupt varactor is typically between -1.2 to -2, as there is a rapid reduction in

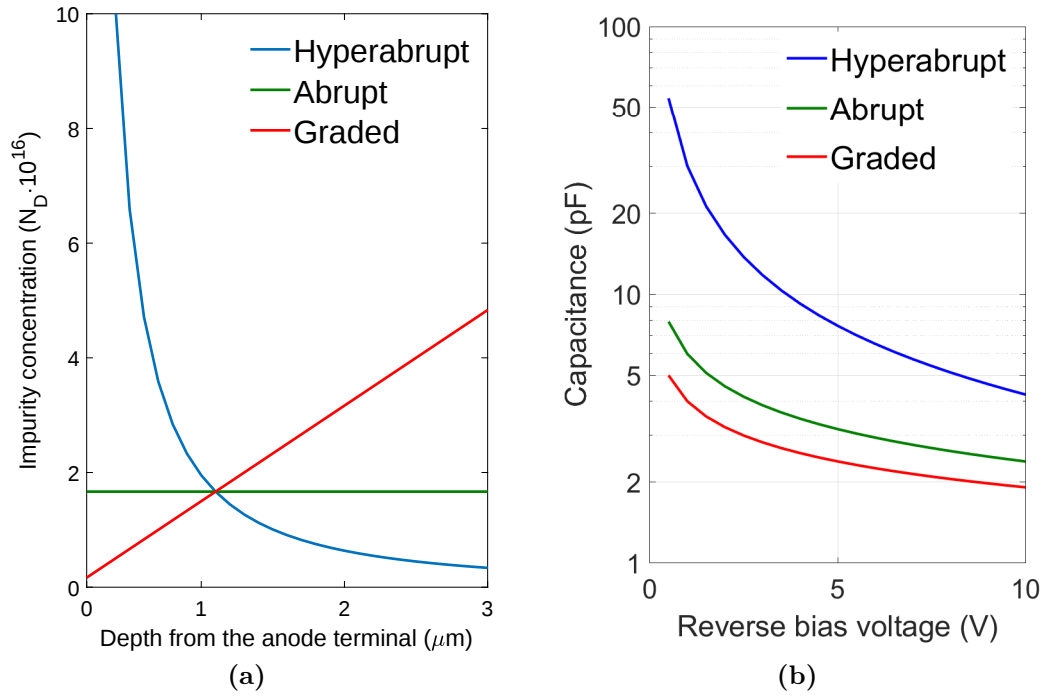


Figure 7.4: Characterization of different grading coefficients on the (a) variation in the doping concentration on the N side of a P⁺/N junction, and (b) capacitance variation with the applied reverse bias voltage V_R .

the charge concentration as the distance from the junction increases. The result is a retrograde dopant profile, with the dopant concentration exhibiting a steep slope near the diffusion region. Complex fabrication techniques with multiple diffusion and/or implantations are required to produce the steep doping profiles needed for hyperabrupt varactors. The nonidealities due to the diffusion process result in varying diffusion coefficients that are dependent on the doping concentration. Hyperabrupt junction varactors have been successfully fabricated in integrated circuits based on silicon (bulk CMOS) [131–133], gallium arsenide [134], and SOI substrates [135], which ensures compatibility with advanced technology nodes.

As shown in Fig. 7.4b, the hyperabrupt varactor offers the greatest variation in

capacitance with an applied reverse bias voltage as compared to an abrupt or graded varactor. For frequency tuning applications, a high capacitance ratio is desired along with high linearity and a large quality factor. Due to limitations in CMOS fabrication processes, it is challenging to simultaneously enhance the tunability, linearity, and quality factor of varactors, which implies tradeoffs between the three parameters are necessary when optimizing the hyperabrupt junction. The quality factor is improved by minimizing the series resistance, which is controlled by the thickness of the epitaxial layer [130]. For decoupling capacitance in digital ICs, the important parameter is to provide a high tunability of the capacitance. Through optimization in the fabrication process, it is feasible to obtain a high capacitance ratio with moderate linearity and a low quality factor [130].

7.2.1 Modeling of a hyperabrupt junction varactor

The electrical model of the hyperabrupt junction varactor is shown in Fig. 7.5. The diode D_S is modeled as a variable junction capacitance $C_J(V)$ and a variable series resistance $R_S(V)$. The series inductance of the diode is negligible [136]. The parasitic impedance of the package and off-chip interconnects connecting the discrete diode includes a series resistance (negligible), parasitic inductance (negligible), and parasitic capacitance C_P . In this work, an on-chip implementation of the hyperabrupt varactor is considered instead of a discrete packaged component. Therefore, the on-chip interconnect is modeled as a parasitic inductance L_P and capacitance C_P as shown in Fig. 7.5.

The junction capacitance is a function of the applied reverse bias voltage and is

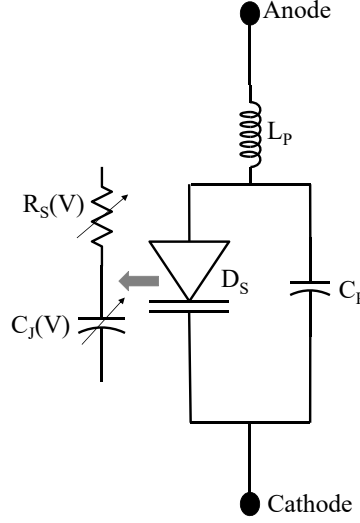


Figure 7.5: Circuit schematic of a discrete hyperabrupt varactor diode [136].

mathematically represented by (7.4). The junction capacitance at zero reverse bias voltage is C_{J0} , and V_J is the junction potential. For silicon, the built-in potential for a P-N junction is 0.47 V under zero bias voltage. The doping profile on the N side of the diode is expressed through the grading coefficient M . Since hyperabrupt varactor diodes exhibit a complex doping profile, the variation of the capacitance with reverse bias voltage is modelled through a curve fitting technique based on (7.4) [136]. The values of V_J and M are, therefore, much higher than that explained through physical properties alone [136]. The three parameters C_{J0} , V_J , and M are defined relative to each other to satisfy the experimentally observed variation in junction capacitance with the applied reverse bias voltage and are obtained through curve fitting the disparate pieces of the measured C-V curve of the varactor [137].

$$C(V) = C_{J0} \cdot \left(1 + \frac{V}{V_J}\right)^{-M} \quad (7.4)$$

The capacitance of a varactor diode is directly proportional to the operating temperature. The empirical formula for the temperature coefficient of capacitance (ppm change in capacitance per °C change in temperature) is given by (7.5) [136], where K is a constant and $\Gamma(V)$ is a function of the reverse bias voltage. Therefore, the greater the tuning ratio of the capacitance, the larger the variation in capacitance with temperature. Most commercially available hyperabrupt varactor diodes exhibit a capacitance variation of $\pm 2\%$ for a temperature range of -40°C to $+80^\circ\text{C}$ [136].

$$T_C(V) = K \cdot \Gamma(V) \cdot 10^6 \quad (7.5)$$

In this chapter, the applicability of hyperabrupt junction varactors as decoupling capacitors for CMOS integrated circuits is explored. The operation of the varactor with an applied DC reverse bias voltage in the sub-volt range (the nominal power supply voltage for a sub 45 nm technology) is considered. A SPICE model of a commercially available hyperabrupt varactor diode with a large capacitance tuning ratio is developed and used to analyze power supply noise suppression. The SPICE model parameters for the SMA1211-001LF diode from *Skyworks Solutions Inc.* are listed in Table 7.2 [138]. The variation of the junction capacitance of the diode with an applied reverse bias voltage is shown in Fig. 7.6. The capacitance changes by 64 pF as the reverse bias voltage is changed from 1 V to 0 V. The model of the varactor is used for SPICE simulations of the different power supply noise scenarios described in Section 7.3. In addition, a Verilog-A model of the hyperabrupt varactor diode is developed to characterize the design space with respect to C_{J0} , V_J , and M .

Table 7.2: SPICE model parameters for the hyperabrupt varactor diode SMVA-001LF [138].

Parameter	Value
Zero-bias junction capacitance, C_{J0}	163 pF
Junction potential, V_J	200 V
Grading coefficient, M	130
Energy gap, E_G	1.11 eV
Saturation current temperature exponent, X_{TI}	3
Saturation current, I_S	10 aA
Series resistance, R_S	0.4Ω
Emission coefficient, N	1
Flicker noise exponent, A_F	1
Forward-bias depletion capacitance coefficient, F_C	0.5
Reverse breakdown voltage, B_V	12 V
Current at reverse breakdown voltage, I_{BV}	1 mA
Recombination current parameter, I_{SR}	0 A
Emission coefficient for ISR, N_R	2
High-injection knee current, I_{KF}	0
Reverse breakdown ideality factor, N_{BV}	1
Low-level reverse breakdown knee current, I_{BVL}	0
Low-level reverse breakdown ideality factor, N_{BVL}	1
Flicker noise frequency exponent, F_{FE}	1
Nominal ambient temperature at which the model parameters are derived, T_{NOM}	27°C

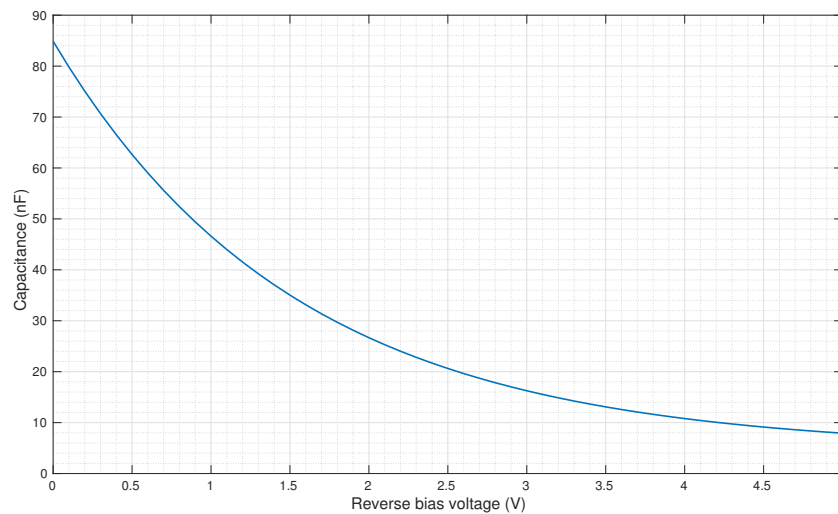


Figure 7.6: SPICE simulation of the capacitance as a function of reverse bias voltage for a hyperabrupt junction varactor based on the Skyworks SMA1211-001LF diode [138].

7.2.2 Power supply noise suppression using hyperabrupt junction diodes

The power supply voltage droop is reduced by increasing the amount of decoupling capacitance as the current delivered through the PDN lags the instantaneous impulse current demanded by the load. If the total capacitance is also a function of the applied voltage, the dependence of the voltage across the capacitor on the charge drawn from the capacitor is no longer linear. The behavior of the hyperabrupt varactor diode is mathematically described in Section 7.1, characterizing the power supply noise suppression the varactor provides when an impulse load current is induced on a clock edge. The corresponding change in the charge of a capacitor for a voltage changing from V_i to V_f is given by (7.6). Similarly, the change in the charge of a reverse biased hyperabrupt varactor diode for a voltage changing from V_i to V_f is given by (7.7).

The capacitance as a function of the applied reverse bias voltage is given by (7.4). Therefore, after substituting for $C(V)$, the change in the charge of a hyperabrupt junction varactor is mathematically expressed as (7.8).

$$\Delta Q_{capacitor} = C \cdot \int_{V_i}^{V_f} dV \quad (7.6)$$

$$\Delta Q_{varactor} = \int_{V_i}^{V_f} C(V) dV \quad (7.7)$$

$$\Delta Q_{varactor} = \frac{C_{J0} \cdot V_J^M}{(1 - M)} \cdot \left((V_f + V_J)^{1-M} - (V_i + V_J)^{1-M} \right) \quad (7.8)$$

For an initial voltage V_i of V_{DD} , the voltage droop across the capacitor when ΔQ charge is required by the load is given by (7.9) and for a hyperabrupt junction varactor diode by (7.10). Using (7.3), the voltage droop of a 48.45 nF on-chip decoupling capacitor is 31 mV for a load sinking 1.55 nC of charge per clock cycle. The characteristics of a hyperabrupt varactor diode are analyzed for varying junction potential V_J and associated grading coefficient M to limit the voltage droop given by (7.10) to less than 31 mV. The characterization of the varactor V_J and M is shown through the bottom surface plot in Fig. 7.8. A junction potential V_J greater than 1 V and a grading coefficient M greater than 2 results in a lower voltage droop than a MIM capacitor with capacitance equal to that offered by a hyperabrupt junction varactor

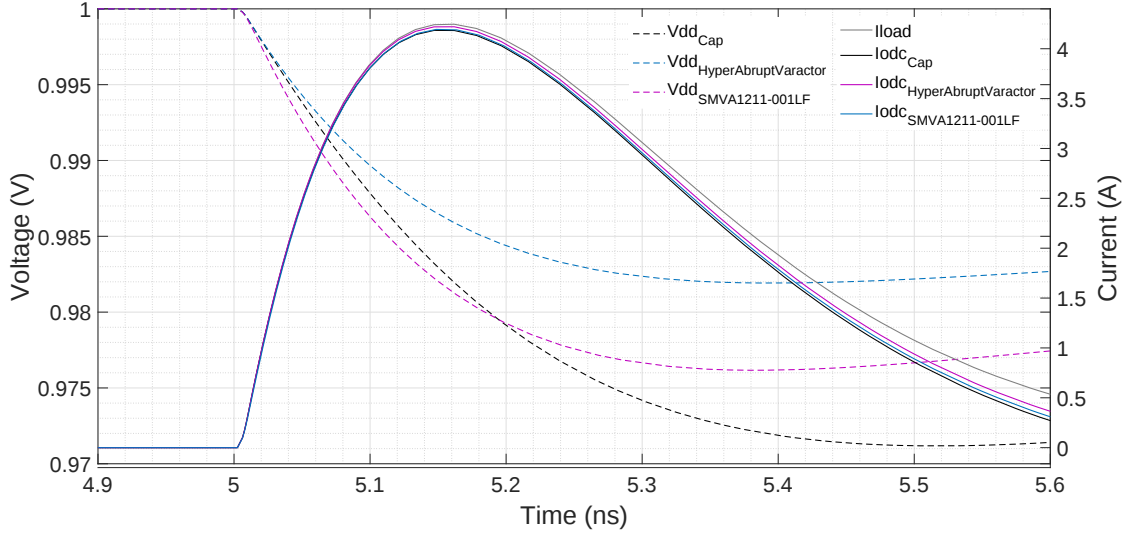


Figure 7.7: Voltage droop due to impulse load current generated on the rising edge of the clock. Three on-chip capacitance scenarios are simulated: ideal capacitor, SMVA-001LF hyperabrupt varactor [138], and custom hyperabrupt varactor ($V_J = 0.7$ V and $M=2$).

diode at a reverse bias voltage of V_{DD} (1 V).

$$V_{droop_cap} = \frac{\Delta Q_{capacitor}}{C} \quad (7.9)$$

$$V_{droop_varactor} = (V_{DD} + V_J) - \left(\frac{\Delta Q_{varactor} \cdot (1 - M)}{C_{J0} \cdot (V_J)^M} + (V_{DD} + V_J)^{(1-M)} \right)^{\frac{1}{1-M}} \quad (7.10)$$

The results indicate that circuit topologies using hyperabrupt varactor diodes to suppress power supply variation due to random load current demand provide benefit. One topology that efficiently suppresses on-chip voltage droop is a series connection of two hyperabrupt varactor diodes, as shown in Fig. 7.9. The series connected varactor diodes are placed as on-chip capacitors close to the load. At steady state, the voltage

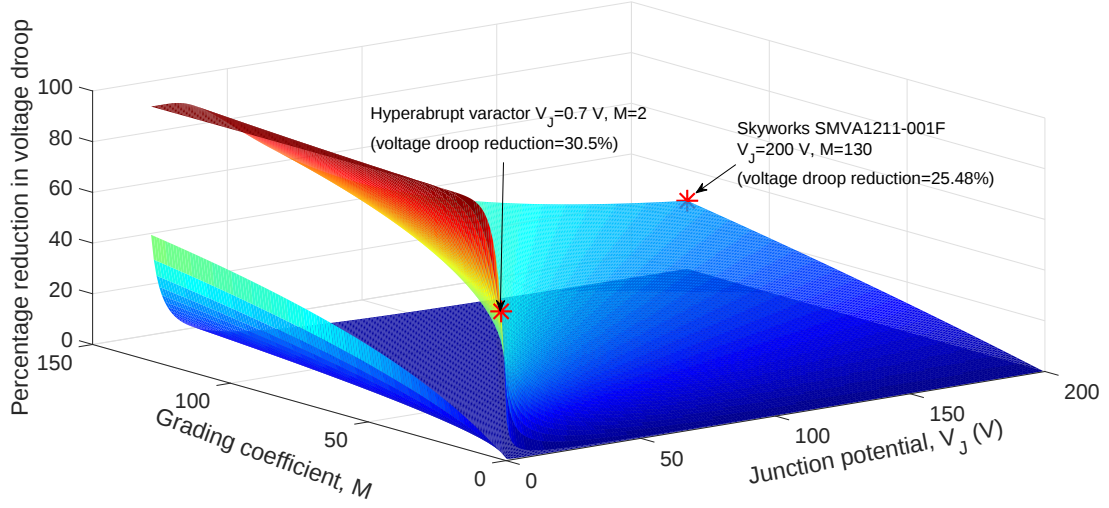


Figure 7.8: Characterization of a hyperabrupt varactor with variation in junction potential and grading coefficient through SPICE simulation. The bottom surface plot is from the characterization of a single varactor, whereas the top surface plot is the result of the characterization of two series connected hyperabrupt junction diodes for the topology shown in Fig. 7.9.

across the two varactors C_1 and C_2 is V_1 and V_2 , respectively. In the event of a charge pulled by the load circuit, the amount of charge provided by both varactors is equal, as given by (7.11). If the final voltage across the varactors C_1 and C_2 is V_x and V_y , respectively, then the voltage droop across the series connected varactors and the load circuit is given by (7.12).

$$\Delta Q_{varactor} = \int_{V_1}^{V_x} C_1(V) dV = \int_{V_2}^{V_y} C_2(V) dV \quad (7.11)$$

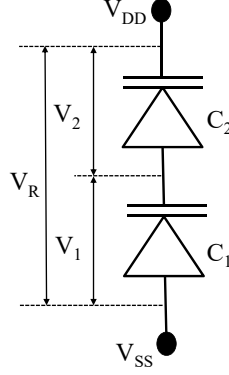


Figure 7.9: Two series connected hyperabrupt junction diodes for on-chip power supply droop mitigation.

$$\begin{aligned}
 V_{droop_varactor} &= V_{DD} - (V_x + V_y) \\
 &= V_{DD} - \left(\left(\frac{\Delta Q_{varactor} \cdot (1-M)}{C_{J0} \cdot (V_J)^M} + (V_1 + V_J)^{(1-M)} \right)^{\frac{1}{1-M}} + \right. \\
 &\quad \left. \left(\frac{\Delta Q_{varactor} \cdot (1-M)}{C_{J0} \cdot (V_J)^M} + (V_2 + V_J)^{(1-M)} \right)^{\frac{1}{1-M}} \right) + 2 \cdot V_J
 \end{aligned} \tag{7.12}$$

The additional reduction in voltage droop with two identical series connected hyperabrupt varactor diodes as compared to a 48.45 nF decoupling capacitor is shown through the top surface plot in Fig. 7.8. The total series capacitance provided by the hyperabrupt junction diodes at a reverse bias voltage of V_{DD} is 48.45 nF. SPICE simulation is run on the switched capacitor load model shown in Fig 7.2 to determine the voltage droop for a range of junction potentials V_J and grading coefficients M . Limited by the current fabrication of the steep doping profiles for hyperabrupt junction diodes, a more than 20% reduction in voltage droop due to clock edge induced impulse current is possible as compared to current MIM decoupling capacitors. The benefits of hyperabrupt junction varactors on noise suppression are shown through the

percentage reduction in voltage droop achieved for a commercially available (SMVA-001LF [138]) and a custom Verilog-A implementation (junction potential V_J of 0.7 V and grading coefficient M of 2) of the hyperabrupt junction diode as marked on the top surface plot shown in Fig. 7.8. Given the substantial reduction achieved in voltage droop, a detailed analysis on power supply noise suppression through series connected hyperabrupt junction varactor diodes is performed in Section 7.3.

7.3 SPICE simulations characterizing noise suppression

The efficacy of series connected hyperabrupt varactor diodes for on-chip noise suppression is characterized by simulating different power supply voltage noise scenarios. The voltage noise on the power supply is a function of the time domain (transient) on-chip current consumption and the frequency domain PDN impedance characteristics. A power delivery network model is constructed with an optimized impedance profile, such that the characteristic impedance of the peaks caused by the parallel resonances of the various components of the PDN are minimized. The transient changes of the current that induce the maximum variation in the power supply voltage on the optimized PDN are simulated for 1) a step up in activity after exiting a power down mode and 2) a resonating event with a frequency close to the resonant frequency due to the on-chip capacitance and package inductance. The suppression of the power supply voltage noise when using hyperabrupt junction diodes is compared with the suppression due to conventional on-chip capacitors.

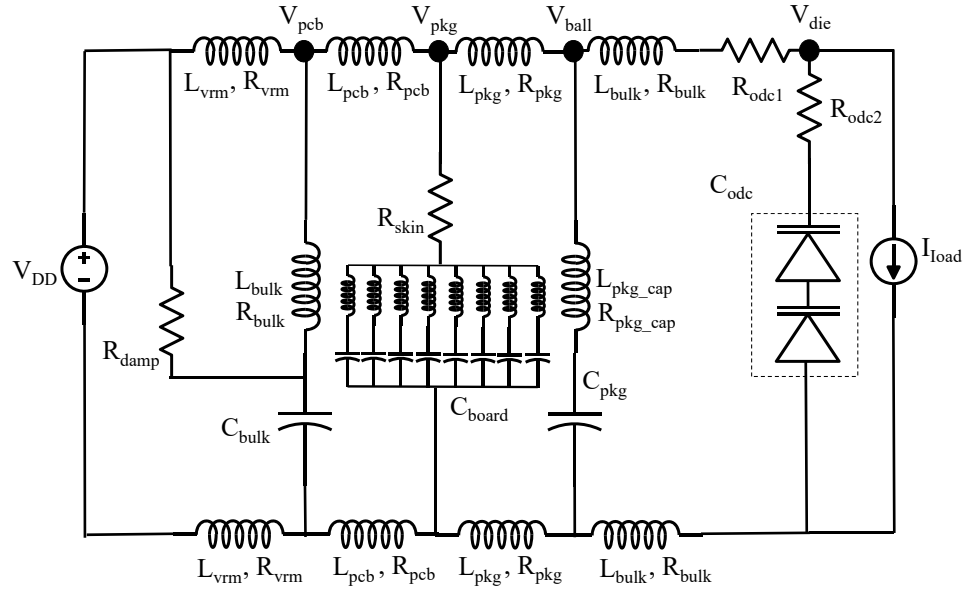


Figure 7.10: Schematic of the optimized power distribution network with distributed multi-layer ceramic capacitors mounted on the PCB and package capacitors.

7.3.1 Construction of the power distribution network

Power delivery to CMOS circuits through off-chip DC-DC regulators results in a PDN with a complex impedance profile. The impedance as a function of frequency exhibits three distinct peaks roughly separated by a decade in frequency. The impedance peaks are due to the resonance between 1) the voltage regulator module (VRM) and the bulk capacitors mounted near the VRM, 2) the loop inductance of the bulk capacitors and the ceramic capacitors mounted on the PCB, and 3) the loop inductance of the ceramic capacitors, package lead inductance, and the on-chip capacitance. The impedance peak due to the resonance between the package inductance and on-chip capacitance has the largest magnitude and is described as the Bandini mountain [129]. The larger the magnitude of the Bandini mountain at a given resonance frequency, the

greater the power supply noise experienced when a transient load current is excited at a frequency close to the resonance frequency of the PDN. Reducing the magnitude of the Bandini mountain is possible by increasing the on-chip capacitance or minimizing the parasitic inductance of the package and the inductance of the multi-layered ceramic capacitors mounted on the PCB.

A PDN model is, therefore, constructed with an impedance profile that minimizes the magnitude of the Bandini mountain for a given maximum on-chip load current. On-package capacitors are added to the model of the PDN. The effective series inductance of the on-package capacitors reduce the C4 bump loop inductance that resonates with the on-chip capacitance, which reduces the height of the Bandini mountain [129]. Ten distributed multi-layer ceramic capacitors are added to suppress the second impedance peak. The impedance peak caused by the VRM is minimized by connecting a shunt resistor, which is present but often ignored when modeling the VRM, between an ideal voltage source and the bulk capacitor. The circuit schematic of the optimized PDN is shown in Fig. 7.10, and the components used to construct the PDN are listed in Table 7.3. The impedance profile of the optimized PDN from the VRM up to the C4 bumps is shown in Fig. 7.11. The low frequency and mid frequency resonance peaks occur at 120 KHz and 10 MHz, respectively. The impedance of the PDN is then determined by considering an on-chip capacitance of 100 nF. The impedance profiles from SPICE simulation for a PDN with an ideal on-chip capacitance and a PDN with an on-chip implementation of two series connected hyperabrupt junction diodes with a total capacitance of 100 nF at a reverse bias voltage of V_{DD} are shown in Fig. 7.12.

Table 7.3: Components of the optimized PDN [129].

Parameter	Value
R_{damp}	10 m Ω
L_{vrm}, R_{vrm}	5 nH, 0.1 m Ω
L_{pcb}, R_{pcb}	48 pH, 0.485 m Ω
L_{bulk}, R_{bulk}	510 pH, 1.365 m Ω
C_{bulk}	132 μ F
L_{pkg}, R_{pkg}	19 pH, 0.123 m Ω
C_{board}	22 nF, 10 μ F, 2.2 μ F (2x), 2 μ F, 440 nF (2x), 100 nF, 4.7 μ F, 940 nF
L_{bump}, R_{bump}	9.5 pH, 0.1 m Ω
L_{pkg_cap}, R_{pkg_cap}	104 pH, 15.72 m Ω
C_{pkg}	440 nF
R_{odc1}, R_{odc2}	1.5 Ω , 1 Ω
C_{odc}	48.45 nF

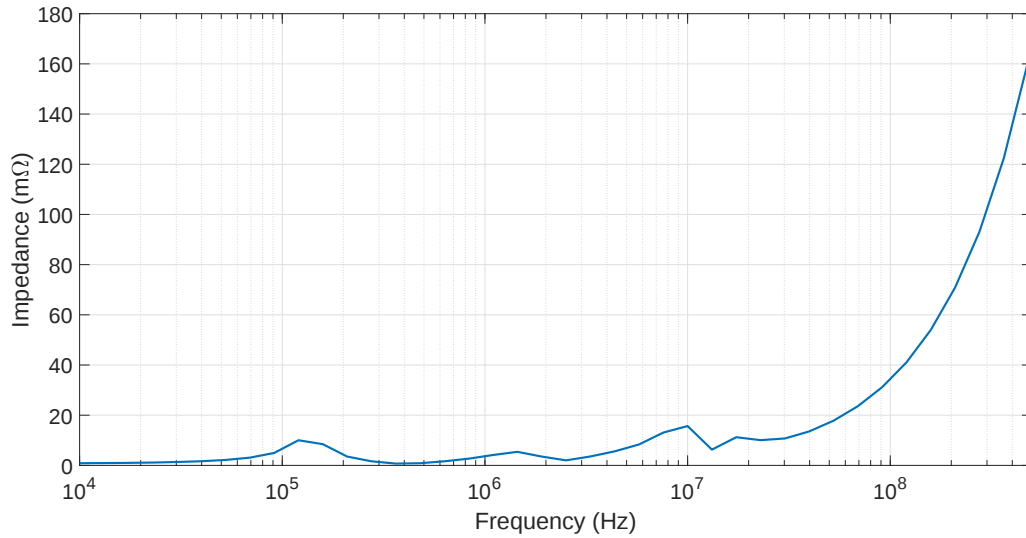


Figure 7.11: Impedance profile of the optimized power distribution network. Due to improved modeling of the voltage regulator, distributed multi-layer ceramic capacitors on the PCB, and on-package capacitors, the impedance peaks at low and mid frequency resonance are minimized [129].

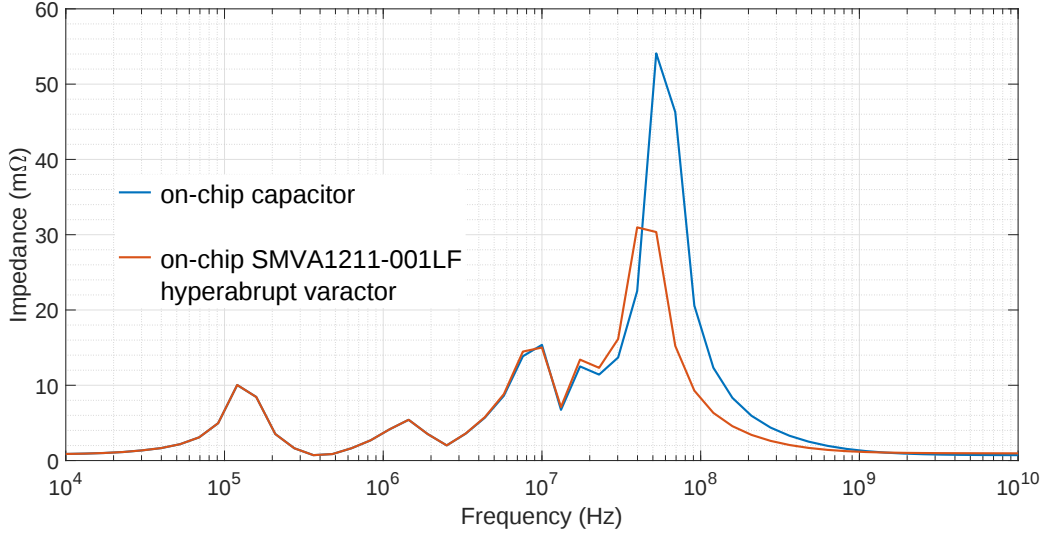


Figure 7.12: Impedance profile of the power distribution network with on-chip capacitance C_{odc} and associated effective series resistance R_{odc} . The impedance profile with an ideal on-chip capacitor (blue) and a series connected SPICE model of a hyperabrupt junction diode (red) [138] are plotted.

7.3.2 Transient current waveforms for analysis of power supply noise

The constructed PDN model is used to characterize the effect that transient current waveforms have on the power supply voltage. The voltage variation on the power supply due to an impulse current on a clock edge has been described in Section 7.1. The other two transient current waveforms of interest are 1) the step and 2) the resonant square wave. SPICE simulations are performed with both waveforms applied to the optimized PDN described in Section 7.3.1. The on-chip capacitance considered for the SPICE simulations is 100 nF. The parasitic resistance contributed by the on-chip capacitor is represented by R_{odc1} and R_{odc2} (refer to Table 7.3). The two transient current waveforms are applied to three circuit scenarios: 1) an ideal 100 nF

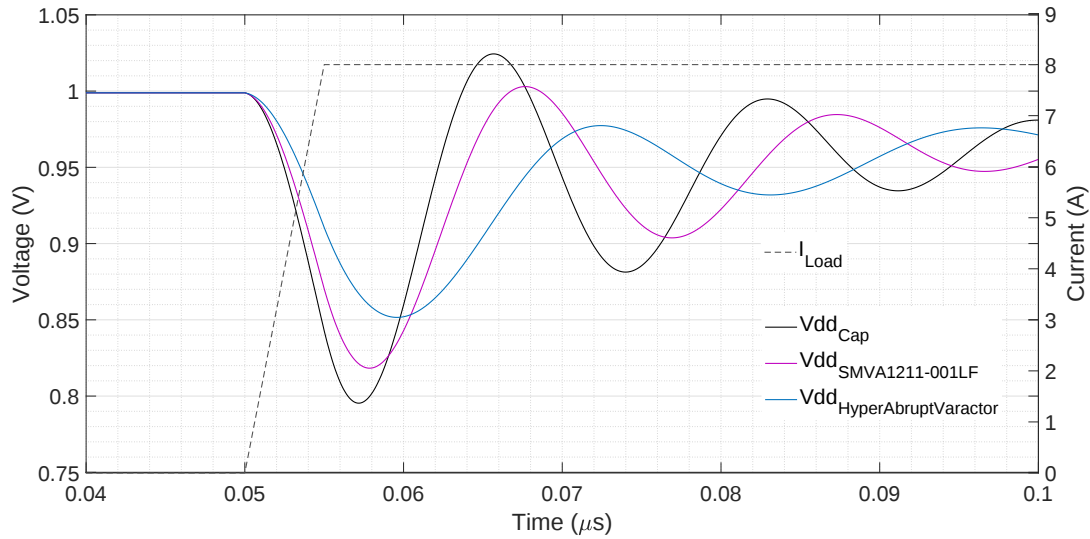


Figure 7.13: Voltage droop due to an 8 A impulse load current.

capacitor, 2) a Verilog-A model of a hyperabrupt junction diode with junction voltage V_J of 0.7 V and grading coefficient M of 2, and 3) a SPICE model of the hyperabrupt junction diode SMVA1211-001LF [138]. Two series connected hyperabrupt junction diodes are considered with a total capacitance of 100 nF at a reverse bias voltage of 1 V (V_{DD}).

Voltage response to current step

The load current I_{load} , as shown in Fig 7.10, is switched from 0 to 8 A in ten clock cycles. A rise time of ten clock cycles is a reasonably aggressive assumption for digital circuits with pipelined architectures where the switching activity increases linearly with each clock cycle [129]. The response of the on-chip supply voltage to a load current emulated as a current step waveform is shown in Fig. 7.13. With an ideal on-chip capacitance of 100 nF, the power supply voltage droops by 205 mV. With the series connected SPICE model parameters of the Skyworks SMVA1211-001LF,

the maximum voltage droop is 181 mV. With the series connected Verilog-A model of the hyperabrupt junction diodes, the maximum voltage droop to an 8 A step current is 149 mV. The series configuration of hyperabrupt varactors as on-chip decoupling capacitors, therefore, offers an 11.7% to 27.3% reduction in the voltage droop as compared to an ideal on-chip capacitance of the same value.

Voltage response to resonating current

The Bandini peak for the optimized PDN with 100 nF of on-chip decoupling capacitance occurs at a resonant frequency of 50 MHz. A repeating load current at a frequency of 50 MHz with 8 A transients is applied as shown in Fig. 7.14. The time domain response of the power supply voltage to the resonating current for the three on-chip capacitance models is also shown in Fig. 7.14. With an ideal on-chip capacitance of 100 nF, the power supply voltage fluctuates by ± 370 mV on either side of the 1 V nominal supply voltage. With the series connected SPICE model of the SMVA1211-001LF, the maximum voltage fluctuation is ± 320 mV. With the series connected Verilog-A model of the hyperabrupt junction diodes, the maximum voltage fluctuation to an 8 A resonating current is ± 160 mV. The high frequency noise caused by the resonance between the on-chip capacitance and the low impedance interconnection between the die and the package is, therefore, suppressed by up to 57% with the series connected hyperabrupt varactors implemented as on-chip decoupling capacitors.

The efficacy of hyperabrupt junction diodes to suppress high frequency noise is demonstrated for a switched capacitor load circuit, which better emulates the load

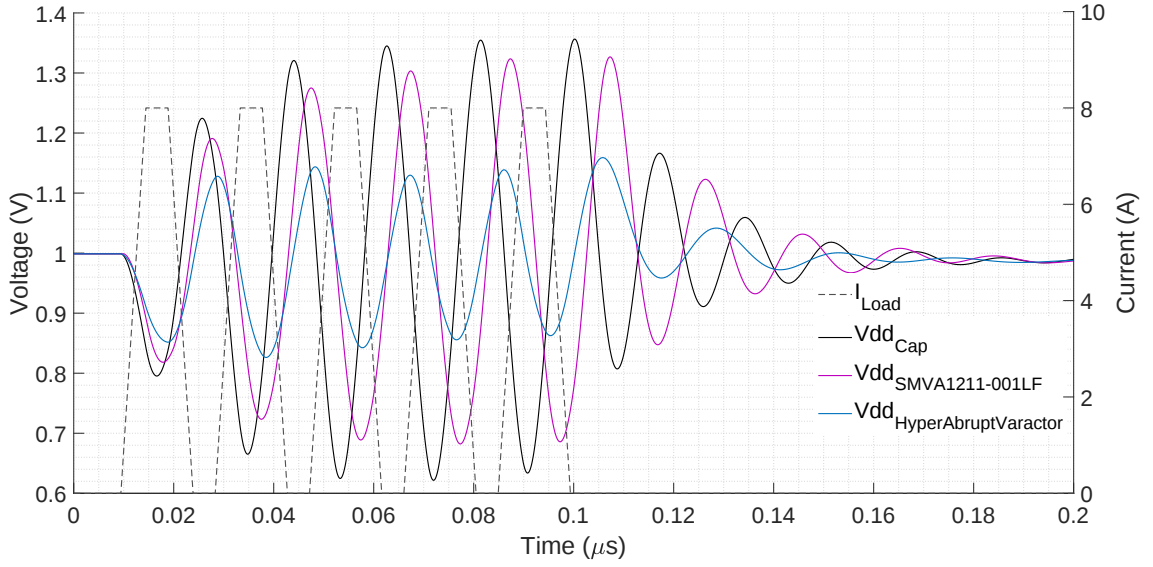


Figure 7.14: Voltage droop on the optimized PDN due to load activity at a resonance frequency of 50 MHz.

current consumption of CMOS circuits (refer to Fig. 7.2). The activity pattern applied to the switched capacitor load circuit is shown in Fig. 7.15(a). The load current I_{load} is used to characterize the current delivered by the three on-chip capacitance configurations; 1) the current I_{Cap} from an ideal capacitor, 2) the current $I_{HyperAbruptVaractor}$ from a series connected Verilog-A model of the hyperabrupt varactor with V_J of 0.7 V and M of 2, and 3) the current $I_{SMVA1211-001LF}$ from a series connected SMVA1211-001LF [138], along with the current I_{PDN} sunk from the PDN for each configuration, are shown, respectively, in Figs. 7.15(c), 7.15(d), and 7.15(e). The 100 mV undershoot and overshoot ($\pm 20\%$ of V_{DD}) when using an on-chip capacitor are suppressed to an 80 mV undershoot and 60 mV overshoot when implementing the SMVA1211-001LF and a 40 mV overshoot and undershoot ($\pm 8\%$ of V_{DD}) with the custom Verilog-A implementation of the hyperabrupt junction diode. The high frequency resonance

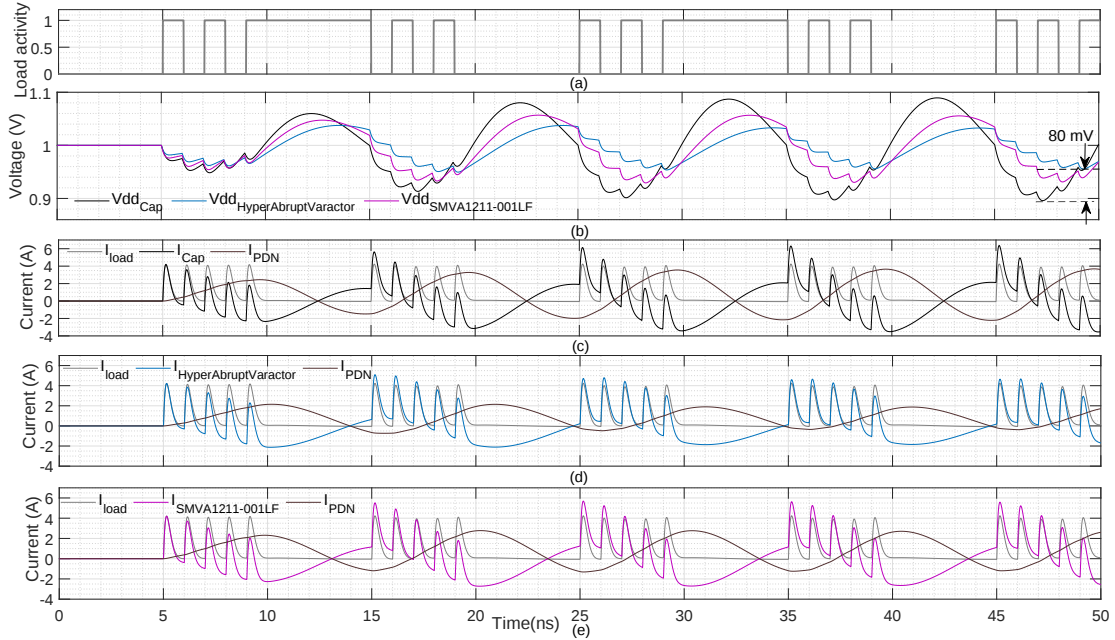


Figure 7.15: On-chip voltage and current response to resonant load activity at 100 MHz to match the impedance peak resonance of the switched capacitor load circuit. The sublets include (a) the resonant wave bit pattern controlling the switched capacitor load circuit, (b) the on-chip voltage noise for the three on-chip capacitor configurations and the current response from the load and on the PDN, and (c), (d), and (e) show, respectively, the three configurations of the on-chip capacitance; ideal, series connected Verilog-A model of hyperabrupt varactor, and SMVA1211-001LF.

noise is, therefore, reduced by up to 60% with the hyperabrupt junction varactors.

7.4 Comparison with existing power supply noise mitigation techniques

In this section, existing on-chip power supply noise mitigation techniques are compared with the proposed implementation using on-chip series connected hyperabrupt junction diodes. Techniques to suppress high frequency noise are considered, which is most critical for on-chip power integrity and broadly consist of power supply noise detection and reaction schemes. Early examples of such techniques applied to on-chip

power distribution in commercial microprocessors are described in [139] and [140]. A power supply droop detection circuit is used to trigger a switched capacitor network, such that the voltage increases through the series combination of the decoupling capacitors. In [141], the on-chip power consumption is monitored and in the event of voltage overshoot or undershoot, the current consumed by the load is controlled to limit the variation in power supply voltage. A PDN conditioner implemented with a large capacitor connected to a high voltage rail is proposed in [142]. The voltage undershoot and overshoot are detected through voltage slope detection circuits. In the event of a voltage droop, a large capacitance charged by a higher voltage rail (greater than 1.5 times V_{DD}) is connected to the V_{DD} rail. Such methods suffer from latency and circuit overhead due to the detection circuits used to monitor power supply variation.

The design constraints listed in Table 7.4 for existing on-chip capacitors are compared with a hyperabrupt junction diode for noise decoupling applications. Note that the leakage current of hyperabrupt junction diodes is inferior to standard decoupling capacitors, where the leakage is a strong function of the operating temperature. The capacitance density is dependent on the P-N junction contact area and is extracted from experimental measurements made in [143]. If the on-chip area is not a constraint, the same amount of voltage noise suppression as obtained from the series connected hyperabrupt varactor diodes is obtained with on-chip decoupling capacitors of twice the capacitance (200 nF). However, increasing the on-chip capacitance results in a degradation in the energy efficiency of the circuit. Two circuit configurations are simulated to illustrate the tradeoff in on-chip area devoted to decoupling capacitance.

The first configuration includes the series connected SMVA1211 varactor diodes with a total capacitance of 100 nF. The second configuration includes 200 nF of on-chip capacitance, an increase of 100 nF to obtain the same amount of power supply noise suppression as the 100 nF varactor diodes. Although the noise suppression achieved for a step load current of 8 A (refer to Fig. 7.13) is the same for the two on-chip configurations, the large on-chip capacitance of 200 nF is charged on every power up cycle of the device. The current required to charge the 200 nF capacitance as compared to the 100 nF varactor diodes is shown in Fig. 7.16(b). The increase in charge consumed on each power cycle or exit from low power state increases the energy consumption of the circuit, which reduces the energy efficiency. For a single power-on event, the cumulative energy consumption to charge the 200 nF on-chip capacitance, normalized with respect to the cumulative energy consumption for the 100 nF series connected hyperabrupt varactor diodes, is shown in Fig. 7.16(c). Despite the larger on-chip area needed to implement series connected hyperabrupt diodes, the energy efficiency of the circuit is greater for the same amount of power supply noise suppression as achieved by deep trench decoupling capacitors.

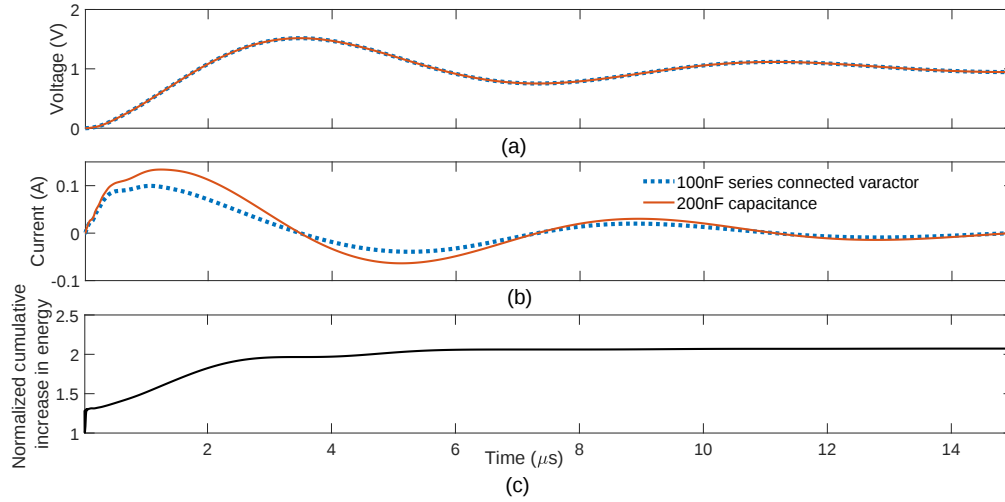
The use of adaptive clocking to suppress power supply noise is a recent and popular technique in high performance processors [125, 144–147]. The technique consists of noise detection and reaction circuits. The reactive component consists of modifying the system clock frequency to negate any timing violations that occur on the critical path of the load circuit when the power supply voltage droops. The voltage droop increases the delay in a digital circuit and a proportional decrease in clock frequency is required to meet the timing constraints. In [146], the magnitude of the voltage

droop is reduced by 17% through adaptive clocking for a step current synchronized across all cores in a 14 nm, 24 core POWER9 processor. In addition to the noise detection latency, adaptive clocking methods incorporate additional delay when the clock frequency is adjusted, including the latency of the modified clock propagating through the clock tree. As described in [125], for a 28 nm dual-core Cortex-A57 cluster operating at 1.1 GHz and 1 V, a clock adaptation latency of less than 2 ns is required to significantly reduce voltage droop without compromising circuit performance.

Implementing the on-chip decoupling capacitors as series connected hyperabrupt junction diodes not only achieves higher suppression in power supply noise as compared to adaptive clocking or other switched decoupling capacitor methods but also achieves noise suppression with minimal impact to latency and performance. In addition, detection of a noise event is not needed, which is a requirement with all existing noise mitigation techniques. Determining the correct triggering instant is a challenge as a late trigger diminishes the maximum possible droop mitigated (and therefore the reduction in voltage margin) and an early trigger potentially impacts performance if the droop is not large. When compared with the complete power supply noise mitigation circuitry (decoupling capacitors, MOS switches, voltage/timing margin detection circuits, adaptive frequency DPLL/PLL circuits), the on-chip series connected implementation of the hyperabrupt junction diode is advantageous to reduce voltage margins while trading-off on-chip area and leakage power consumption.

Table 7.4: Characterization of leakage current, capacitor density, and temperature stability for different families of capacitors.

Capacitor type	Leakage current ($\text{A}/\mu\text{m}^2$)	Capacitor density ($\text{fF}/\mu\text{m}^2$)	Temperature stability ($\text{ppm}/^\circ\text{C}$)
CMOS Decap (65 nm process) [123]	10^{-9}	6.6	50 to 100
2.5 V NMOS thick oxide decap (65 nm process) [123]	10^{-9}	3.7	30 to 100
MIM decap	10^{-17} [148]	0.38 to 1.1 [148, 149]	35 [150]
Deep trench	$<10^{-15}$ for MIMIMIM deep trench [151]	158 (65 nm process), >400 in 32 nm and below [151, 152]	not provided
Hyperabrupt junction diode	10^{-10} to 10^{-7} at $V_R = 1 \text{ V}$, $T = 25^\circ\text{C}$ [136]	400 (65 nm process) at $V_R = 1 \text{ V}$, $T = 25^\circ\text{C}$ [143]	100 to 300 [136]

**Figure 7.16:** The energy overhead of introducing additional on-chip decoupling capacitance to obtain the same amount of power supply noise suppression as series connected hyperabrupt junction diodes. Shown in the sub-figures are (a) the on-chip voltage variation as the decoupling capacitors are charged at power up, (b) the current sunk from the DC-DC regulator at power up through the optimized PDN by the on-chip decoupling capacitance, and (c) the cumulative energy consumed during power up for the overcompensated on-chip capacitance of 200 nF normalized to the cumulative energy consumed with 100 nF of series connected hyperabrupt varactor diodes.

7.5 Summary

Varactors with high capacitance tuning ratios such as the hyperabrupt junction diode are proposed for on-chip power supply noise suppression for noise sensitive digital blocks. The increase in capacitance as the voltage across the varactor drops is exploited to reduce the dependence of the voltage across the varactor terminals

on the charge stored or released from the varactor. For the same amount of charge drawn, the voltage drop across series connected hyperabrupt junction diodes is less than a MIM or deep trench capacitor with the same capacitance. The reduction in the variation of the power supply voltage with on-chip hyperabrupt junction diodes is simulated with a current model that includes an optimized power delivery network for high frequency resonance, a step current, and a clock edge induced current. The three current waveforms are used to characterize the stability of the PDN. The proposed circuit technique with a commercially available hyperabrupt junction diode and with a custom hyperabrupt junction diode implemented with a junction voltage V_J of 0.7 V and grading coefficient M of 2, show, respectively, up to 40% and 60% reduction in the peak magnitude of high frequency power supply voltage noise as compared to current on-chip decoupling capacitors. The efficacy of the proposed noise suppression technique is demonstrated for the same amount of on-chip decoupling capacitance. In addition, there is no performance penalty or additional circuitry needed as compared to existing power supply noise mitigation techniques found in commercial processors and literature.

Chapter 8: Run-time power supply voltage detection and clamping mechanism for 3-D ICs

"[3-D integration] technology provides tremendous flexibility as designers seek to "mix and match" IP blocks with various memory and I/O elements in new device form factors."

– Intel newsroom release on 3D logic chip "Foveros", December 2018.

The high power consumption and signal delay due to global interconnects in nano-scale technologies has produced research in novel integration technologies. One novel research area is through silicon via (TSV) based 3-D IC technology, where each layer or stratum is fabricated separately and subsequently vertically integrated [153]. TSV based 3-D integrated circuits permit the integration of heterogeneous technologies with CMOS. The integrated system may include RF, analog, micro-/nano-electromechanical systems (MEMS/NEMS) as well as emerging technologies such as nano-FET and graphene-based device planes.

With 3-D ICs, it is possible that the fabrication of disparate strata and the final integration takes place in separate manufacturing facilities. The packaging manufacturer performing the final bonding does not necessarily need the technology dependent parameters, including the voltage levels and frequency of operation for devices and IOs, of each stratum. The transparency to technology information opens the possibility for off-the-shelf integration, where dies from different foundries are bonded together by the packaging manufacturer to form a heterogeneous 3-D IC. There are

two criteria to facilitate a “plug-and-play” approach to 3-D IC integration. First, guidelines to standardize the interface circuit properties (location and electrical characteristics of IO ports and ESD protection) for each device plane must be developed and implemented. Second, the global power and clock generation circuits must meet the requirements of each stratum in the 3-D IC stack.

In this chapter circuit techniques to address the global power generation and distribution of a 3-D IC are provided. The goal is to develop a power delivery system that senses the power supply voltage needed by each voltage domain V_{DD_DP} in each device plane at run time. A dedicated power plane is implemented to deliver power to all voltage planes in a 3-D stack. The power plane is capable of generating a range of programmable voltage levels, ideally from 0.5 V to 5 V, to meet the power supply voltage requirements of various technology nodes. Depending upon the 3-D IC stack configuration and power budget, single or multiple power planes are interspersed between device planes. The goal is to provide power to a heterogeneous device configuration that includes CMOS, MEMS, or RFICs operating on disparate substrates not limited to Silicon or III-V technologies. A 3-D IC that consists of digital, mixed signal, and IO circuits requires different voltages. The power supply voltage requirement of each voltage domain on each device plane is auto-detected by novel circuits found on the dedicated power plane and the heterogeneous device planes.

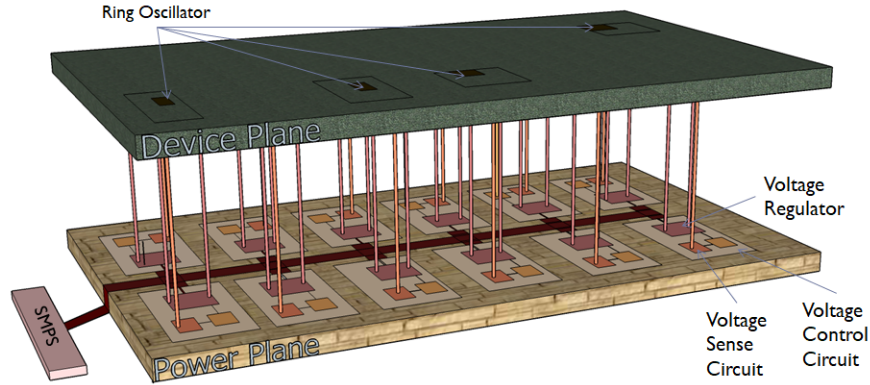


Figure 8.1: Developed voltage sense and power delivery system for 3-D ICs. A 3-D stack with device planes containing multiple voltage domains are served by a power plane consisting of multiple power modules. The ring oscillator placed in each voltage domain provides the controlling frequency to the voltage sense circuit in the power module. Two circuit techniques for the voltage sense and control mechanism are developed that provide a precise reference voltage to the on-chip voltage regulator.

8.1 Power supply voltage detection and clamping technique

The voltage sense circuit is implemented using a ring oscillator. Each voltage domain within each device plane includes a ring oscillator capable of generating an output frequency F_{out} of 1 GHz when a control voltage equal to the power supply voltage V_{DD_DP} of the domain is applied. A 1 GHz frequency is an arbitrarily chosen value for the proposed circuit and is adjusted based on set standards. The goal is to select a common frequency of operation for all ring oscillators placed on different device planes in a given 3-D IC stack. The ring oscillator is the only required overhead for each voltage domain. The remaining components of the voltage sense circuit are placed on the dedicated power plane. The F_{out} generated by the ring oscillator propagates through a TSV to the power plane. Two disparate circuits are developed

for the power plane [154–156]. The block diagram of the ring oscillator based voltage detection and clamping circuit is shown in Fig. 8.1. The construction and operation of all the components included in the voltage sense, control, and power delivery circuit are described in subsequent sections.

8.2 Ring oscillator circuit for each voltage domain

Ring oscillators (RO) are widely used as voltage controlled oscillators (VCO) in high performance integrated circuits as the fundamental block for frequency synthesizers, clock recovery circuits, and clock distribution networks. The application is not limited to VCOs, as ring oscillators are used as on-chip thermal sensors [157], [158] and test structures to measure process variability [159]. The versatility of a ring oscillator is attributed to a simple CMOS implementation with no passive components. A simple RO implementation minimizes the occupied silicon area and stabilizes the output frequency in the presence of process, voltage, and temperature variation (PVT) and the applied control voltage.

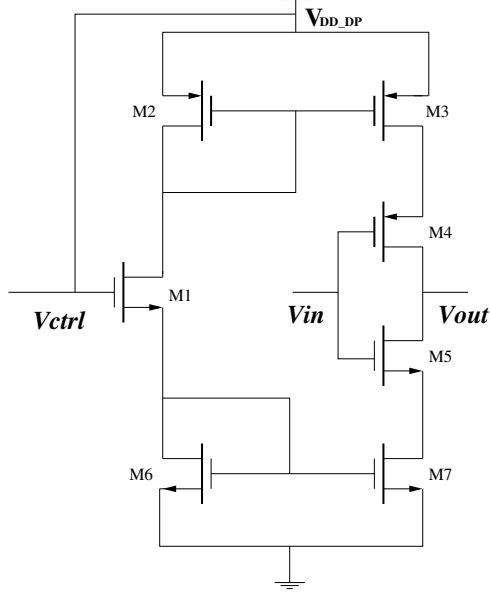


Figure 8.2: Current starved with output switching inverter.

Several ring oscillator circuit topologies are characterized for the detection of the targeted V_{DD_DP} of a given voltage domain. The selected ring oscillator circuit topology must provide a large frequency range for the desired application with minimum variation in frequency due to sensitivity to PVT. A current starved RO with an output switching inverter (refer to Fig. 8.2) provides the best frequency stability in terms of temperature variation (less than 2%), a low phase noise of 0.06 radians, and frequency sensitivity to power supply variation of less than 10% [160], [161]. An addition based current source [162], which maintains a constant total current across process variations, is used as a header and footer circuit for the CMOS inverters cascaded to form a ring oscillator in [163]. Despite the stability of the output frequency with process variation and device mismatch, the ring oscillator topology in [163] is not suitable for the problem addressed in this chapter due to:

- 1) The control voltage V_{ctrl} to the gate of transistor M1 must be higher than the

threshold voltage V_{TH} of the transistor and lower than V_{DD} . For nanoscale technologies, where the difference between V_{TH} and V_{DD} is not large, the operating range of V_{ctrl} is limited, which implies a narrow voltage range for the controlling current feeding the current starved inverters. The addition based current source implemented in [162] is shown in Fig. 8.3.

2) A single voltage that serves as both V_{DD} and the control voltage to the inverter stages is preferred, however, the current circuit includes two separate voltages. The current starved topology with an output switching inverter, is therefore, selected as the building block for the ring oscillator implementation. The control voltage V_{ctrl} is applied to both the control transistor M1 and V_{DD_DP} as shown in Fig. 8.2.

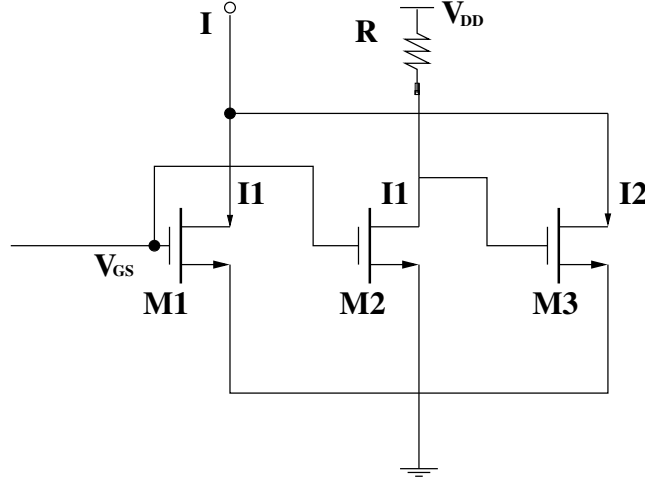
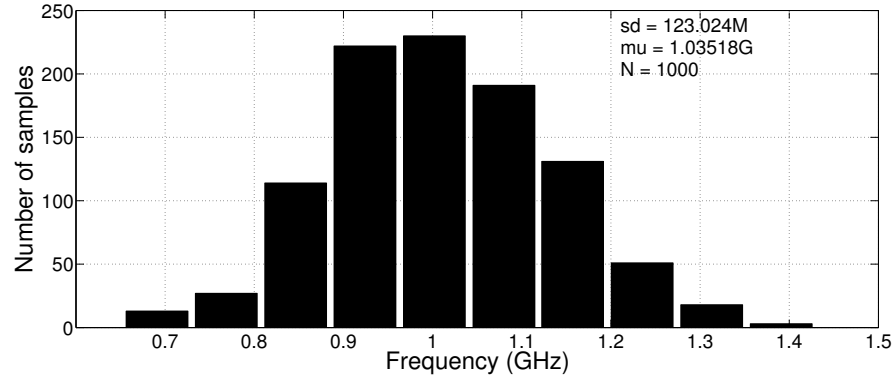


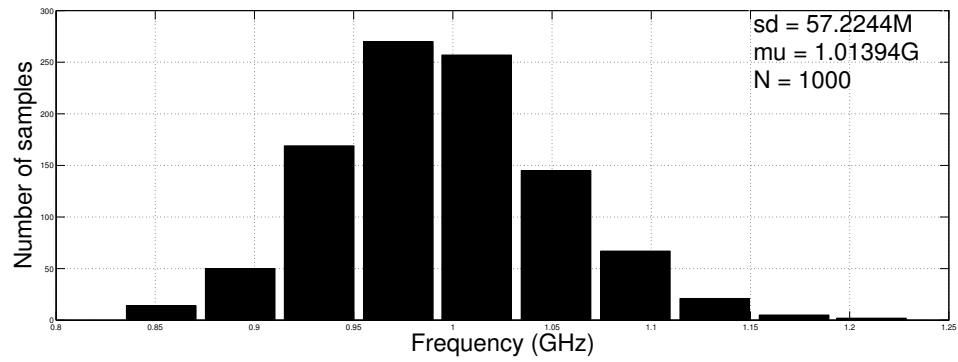
Figure 8.3: Addition based current source used as footer circuit in [162].

The RO circuit is implemented using a 22 nm high performance (HP) predictive technology model (PTM) model [66] with a V_{DD_DP} of 0.8 V. An output frequency of 1 GHz is achieved with three current starved inverting stages each with transistor W/L ratios between 8 to 10. None of the transistors are minimum sized to reduce the impact of line-edge roughness and random dopant fluctuations that cause significant

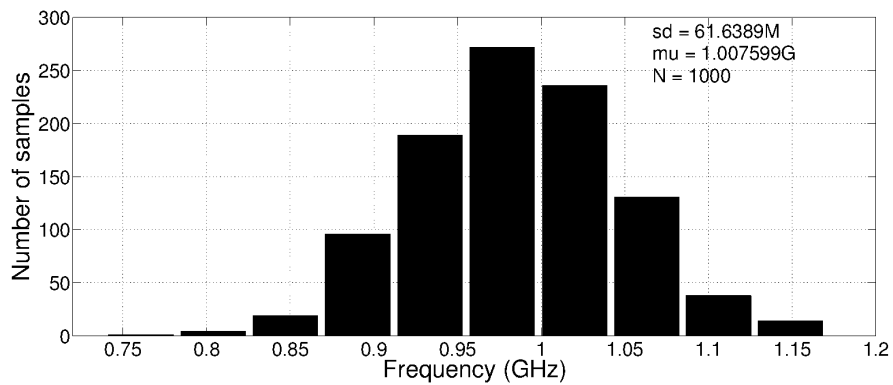
variation in the V_{TH} in sub-nanometer technology nodes [164]. The V_{TH} , low-field mobility, and effective channel length are the three main parameters impacted by lithographic variation, stress, and doping concentration in strained silicon technology. The impact of process variation on the output frequency of the ring oscillator is evaluated with Monte Carlo analysis. A typical corner is simulated for statistical variation of V_{TH} , channel length, and mobility considering both process variation and device mismatch. The variation in the ring oscillator frequency variation with V_{TH} , channel length, and MOSFET low-field mobility is shown, respectively, in Figs. 8.4a, 8.4b, and 8.4c. The corresponding ratio of the variance to the mean (σ/μ) for V_{TH} , effective channel length, and low-field mobility is, respectively, 11.9%, 5.64%, and 6.12%. The designed ring oscillator therefore exhibits moderate deviation in output frequency with process variation. The power module on the power plane that detects the ring oscillator frequency is designed to compensate for deviation in frequency from 1 GHz due to PVT variation.



(a)



(b)



(c)

Figure 8.4: Statistical variation of ring oscillator output frequency for 22 nm HP PTM for (a) V_{TH} , (b) effective channel length, and (c) low-field mobility.

8.3 Method I for power plane construction: Frequency to voltage converter topology

A frequency to voltage converter (FVC) placed on the power plane generates a voltage V_{fvc} inversely proportional to F_{out} . The non-linear voltage from the FVC controls a current source that provides the control voltage to the ring oscillator and the power supply voltage to the voltage domain served by the ring oscillator. Dedicated power modules comprising of the FVC and a dependent current source for each voltage domain facilitate point of load power delivery which offers three distinct advantages [165]:

1. Reduced noise due to a reduction of the parasitic impedance of the power distribution network as the voltage source is closer to the load circuit,
2. Supply of different voltages to heterogeneous circuits, and
3. Finer granularity for voltage control.

The schematic of the proposed FVC based voltage detection and clamping technique is shown in Fig. 8.5.

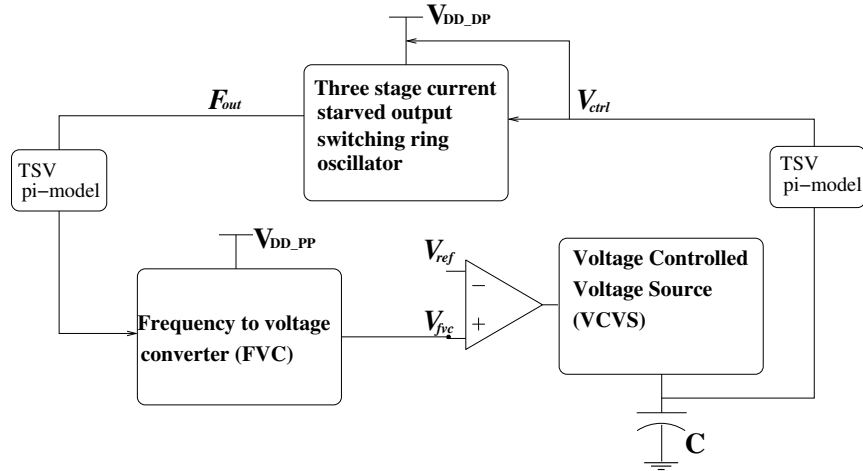


Figure 8.5: Block level schematic of the voltage detection and power delivery circuit for a heterogeneous 3-D IC.

8.3.1 Frequency to voltage converter circuit

The output of the ring oscillator on the device plane is connected through TSVs to a FVC circuit placed on the dedicated power plane. With careful floor planning and the development of 3-D IC interface circuit standards, the vertical alignment of the ring oscillators and the FVC circuits on the respective device and power planes ensures minimum electrical degradation of the frequency signal in terms of induced jitter and propagation delay. The circuit chosen to implement the FVC is described in [166]. The FVC operates based on charge redistribution between two switching capacitors. The generated output voltage is linearly related to the input frequency and does not require filtering to remove AC ripples, which is a common limitation with other FVC implementations using low pass filtering techniques or digital counters. The circuit implementation is also area efficient, as a limited number of components are required.

The FVC circuit is implemented using a 45 nm low power (LP) PTM model [66] with the power supply voltage V_{DD_PP} set to 1 V. The operation of the FVC is shown in Fig. 8.6 [166]. The current source is set to 20 μ A. The capacitors C_1 and C_2 are each 20 fF and are implemented as MOSCAPS to reduce the occupied area. The voltage across the capacitors is inversely proportional to the frequency of the input signal and is approximately 380 mV when the input frequency is 1 GHz.

8.3.2 Implementation of the voltage controlled voltage source

The output voltage from the FVC controls a dependent current source implemented using the thick oxide PTM model available through the NCSU 45 nm PDK [167]. The dependent current source is connected to a capacitor and provides a ramp voltage when the controlling voltage is constant. The current source provides a gain

sufficient to generate stable output voltages up to 3 V, which is sufficient to serve analog and IO devices implemented in sub-nanometer technologies. Generating output voltages up to 5 V is possible by selecting an older technology node (greater than 65 nm) to implement the current source on the dedicated power plane. By supplying voltages from 0.5 V to 5 V, a wide range of disparate technologies are integrated to form a 3-D system. The voltage output from the ramp generator circuit acts as the controlling voltage V_{ctrl} for the ring oscillator placed on the device plane.

The ring oscillator generates distinct frequencies when the input control voltage falls between $V_{DD_DP}/2$ and V_{DD_DP} . The FVC responds to the frequencies by providing an output voltage inversely proportional to the input frequencies. The current source generates a proportional current based on output voltage from the FVC. The current is converted to a non-linear control voltage that ensures a faster transition to the desired power supply voltage (V_{DD_DP}) of a given voltage domain.

8.3.3 Simulation results

The three circuit sub-blocks, which include the 22 nm three stage current starved output switching ring oscillator circuit on the device plane (power supply voltage V_{DD_DP} of 0.8 V), the 45 nm FVC circuit (power supply voltage V_{DD_PP} of 1 V), and the 45 nm voltage controlled voltage source (VCVS), are combined and simulated. The simulation also includes the impedance of the TSVs. The topology of the circuit used for simulation is shown in Fig. 8.5 and the transient response is shown in Fig. 8.7. Assuming a vertical alignment between the power module on the power plane and the ring oscillator on the device plane, the interconnect length is negligible and

is not considered in the simulation.

The control voltage from the VCVS and the output signal of the ring oscillator propagate through TSVs. The electrical impedance of the TSVs [168], [169] is represented as an equivalent RC pi-model. The values of the DC resistance (505.8 m Ω), 1 GHz resistance (570.72 m Ω), and capacitance (8.7 fF) of a single TSV are computed from the closed form expressions developed in [170]. The pi-model represents a tungsten filled TSV with a length of 16 μm , diameter of 1.5 μm , and dielectric thickness of 0.25 μm . Each RC pi-model shown in Fig. 8.5 represents two TSVs in parallel.

The transient response of the 22 nm ring oscillator (F_{out_22nm}), the FVC output voltage (V_{fvc}), and the control voltage generated by the VCVS block (V_{ctrl_22nm}) is shown in Fig. 8.7. The ring oscillator is the only current load served by the power module for the transient simulation. The VCVS transient simulation, therefore, excludes the impact of load transients and resistive losses due to the power delivery network and additional load circuits. The desired power supply voltage of 0.8 V is reached in less than 80 ns. The proposed voltage detection and power delivery circuit is further tested to provide the power supply voltage to a 45 nm device plane. A three stage current starved ring oscillator is implemented using a 45 nm LP PTM model (V_{DD_DP} of 1 V). The VCVS is able to reliably generate a control voltage of 1 V (V_{ctrl_45nm}) in less than 100 ns, as shown in Fig. 8.7(c).

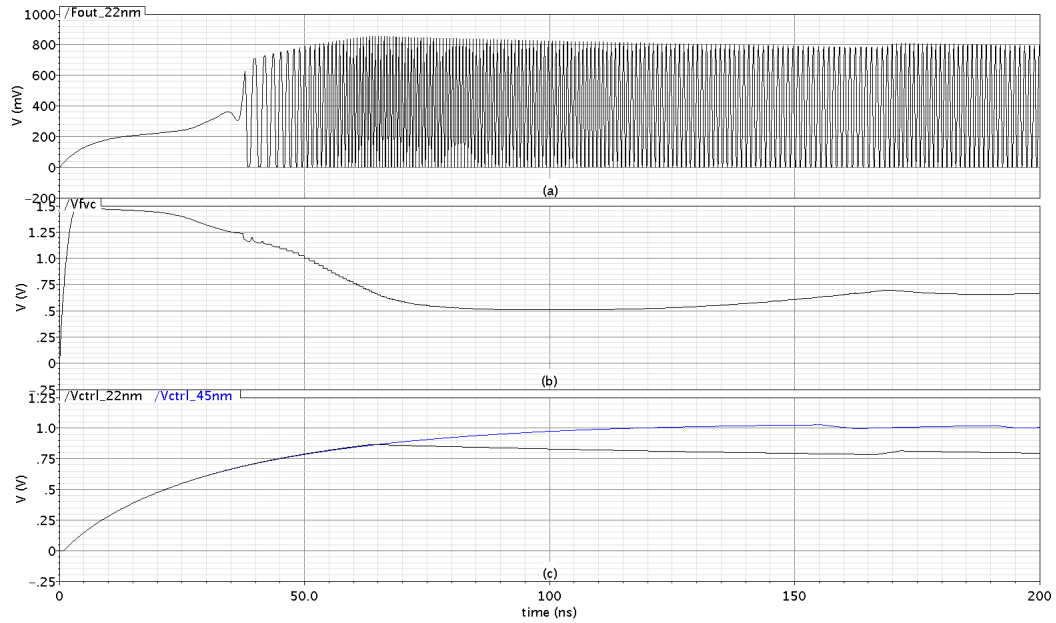


Figure 8.7: Transient response of the (a) ring oscillator placed in a 22 nm device plane, (b) FVC, and (c) control voltage generated from the VCVS for both a 22 nm (V_{DD_DP} of 0.8 V) and 45 nm device plane (V_{DD_DP} of 1 V).

8.4 Method II for power plane construction: Frequency comparator topology

In a second implementation of the power module that serves a voltage domain, a clock divider, a frequency comparator, a voltage ramp generator, a voltage peak detector, and a voltage regulator are used. The schematic for the voltage detection and clamping circuit based on a frequency comparator topology is shown in Fig. 8.8.

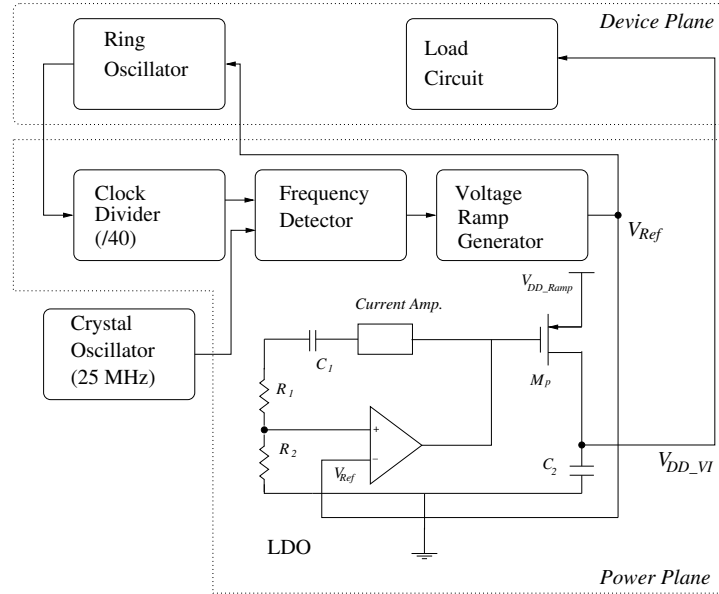


Figure 8.8: Block level view of the voltage detection and clamping circuit using a frequency comparator based topology.

8.4.1 Divide by 40 clock divider

The generated signal from the ring oscillator that propagates through a TSV into the power plane is down converted to a 25 MHz signal. A clock divider consisting of a divide-by-8 cascaded with a divide-by-5 circuit block is designed for the down conversion of the RO output. A non-binary CMOS divider topology [171] which provides a non binary dividing ratio without the use of a decoder, is selected to implement the clock divider circuits. There is no degradation in the signal propagation delay as compared to decoder based non-binary dividers.

8.4.2 Frequency comparator

A frequency comparator compares the output of the clock divider with a 25 MHz clock source from an off-chip crystal oscillator. The frequency comparator is implemented using four D-FFs as shown in Fig. 8.9(a). The frequency comparator is

chosen over a phase frequency detector (PFD) as the PFD generates UP and DOWN signals (due to D-FF clock-to-Q delay) even when the inputs are phase locked [172]. In addition, the voltage detection circuit does not require that the down converted ring oscillator frequency is matched in phase with the 25 MHz reference signal.

8.4.3 Voltage ramp generator

A voltage ramp generator circuit (RGC) includes a constant current source charging a capacitor C_1 and a switch S_1 that controls the duration of the ramp voltage, as shown in Fig. 8.9(b). The UP and inverted DOWN signal are logically ANDed ($UP \bullet \overline{DOWN}$) to generate the control signal for S_1 . When the 3-D IC is first powered on, the ramp generator provides an initial voltage to the ring oscillator, and the UP signal from the frequency detector is high as the down converted frequency from the ring oscillator is less than 25 MHz. The frequency increases as the output voltage from the ramp generator increases. When the ring oscillator reaches the 1 GHz target frequency, the UP signal is de-asserted and any further increase in frequency asserts the DOWN signal. The switch S_1 is in the open state, preventing any further increase of charge (and therefore voltage) on C_1 . The voltage V_{RGC} on C_1 corresponds to the voltage domain V_{DD_VI} in which the ring oscillator is placed. The discharging of the capacitor C_1 through the load circuit causes the ring oscillator frequency to drop below 1 GHz, which re-asserts the UP signal and places S_1 in the closed state. The UP signal continues to periodically toggle ensuring that the voltage V_{RGC} is maintained at the desired level.

8.4.4 Voltage peak detector

The variations in V_{RGC} due to the charging and discharging of the capacitor C_1 are filtered using a voltage peak detector circuit (shown in Fig. 8.9c). The peak detector circuit consists of a PMOS M_1 that controls the current charging the capacitor C_2 . A voltage comparator compares the output voltage V_{Ref} across the capacitor C_2 with V_{RGC} . The output voltage from the comparator biases M_1 . V_{Ref} follows the positive transition of V_{RGC} and at steady state equals the maximum value of V_{RGC} . The peak detection circuit, therefore, provides a steady voltage reference with less than 1% ripple voltage variation from the targeted power supply voltage of the device plane (V_{DD_VI}). V_{Ref} is used as the reference voltage to the voltage regulator serving the load circuit in the device plane. On-chip voltage regulator topologies like the LDO and buck converter are suitable for integration with the proposed voltage detection and clamping circuit. The stable reference voltage provided by the proposed circuit ensures superior line regulation offered by the voltage regulator.

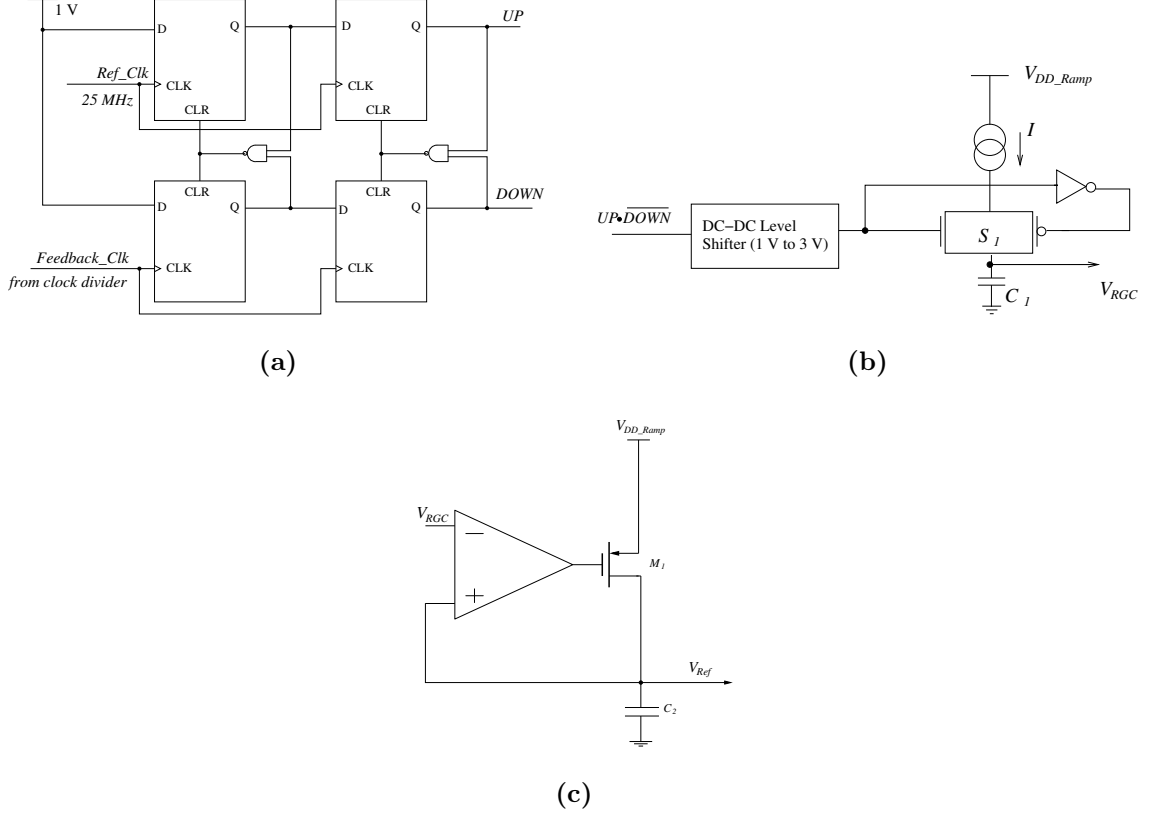


Figure 8.9: Circuit schematic of (a) frequency detector, (b) voltage ramp generator, and (c) peak detector.

8.4.5 Simulated Circuit Results

Two device planes are simulated using the 45 nm low-performance (LP) (V_{DD_VI} of 1 V) and 22 nm high-performance (HP) (V_{DD_VI} of 0.8 V) PTM models [66]. The designed ring oscillators for each device plane include three inverter stages in a current starved output switching configuration.

The components of the power module are implemented with the 45 nm PTM models. The divide-by-40 clock divider and frequency detector circuits are simulated using the 45 nm LP PTM model (V_{DD_PP} of 1 V). The voltage ramp generator and peak detector circuits are implemented using the 45 nm thick oxide (V_{DD_Ramp} of

3 V) PTM model. A DC-DC level shifter is used to convert the UP and DOWN signals from a V_{DD_PP} of 1 V to a V_{DD_Ramp} of 3 V. The slope of the voltage ramp signal is deliberately kept low (0 V to 3 V in 2 μ s) to ensure stable operation. The minimum voltage reliably detected by the power module is 0.7 V. The maximum voltage provided by the ramp generator is 2.5 V. SPICE circuit simulations indicate a maximum variation of 1% in the reference voltage V_{Ref} provided to the on-chip voltage regulator for V_{DD_VI} of less than 1 V. The stability of the reference voltages is comparable to the stability of off-chip buck converters, where approximately 1% variation is currently achieved [173].

The simulation results for detecting and clamping the power supply voltage for a voltage domain in a 22 nm device plane are shown in Fig. 8.10. V_{Ref} reaches a V_{DD_VI} of 0.8 V in 370 ns. The detection and clamping circuits on the 45 nm device plane require 420 ns for V_{Ref} to reach the target voltage V_{DD_VI} of 1 V. The proposed circuit, therefore, offers a fast transition to the desired voltage level at startup and is suitable for integration with on-chip voltage regulators.

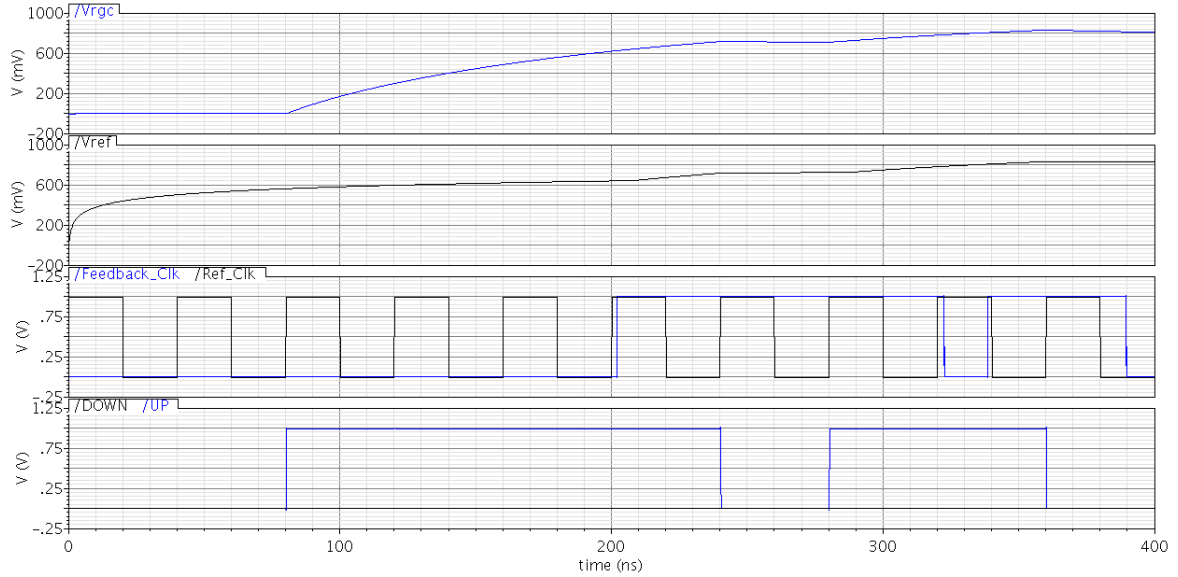


Figure 8.10: SPICE analysis of the power supply voltage detection and clamping circuit for a 0.8 V domain in a 22 nm technology.

8.5 Summary

Two mechanisms to detect the power supply voltage of a given voltage domain in a 3-D IC are implemented by placing a ring oscillator in each voltage domain located on disparate device planes. The chosen circuit topology for the ring oscillator exhibits acceptable deviations due to PVT variations in a 22 nm technology. A σ/μ ratio of 11.9%, 5.64%, and 6.12% is determined for, respectively, threshold voltage variation, channel length variation, and low-field mobility. The power requirement of each voltage domain is served by a single power plane that includes multiple point of load power modules.

In the first circuit implementation, each power module consists of a voltage controlled current source regulated by a frequency to voltage converter. The power supply voltage detection and delivery mechanism is demonstrated by simulating the

device plane and power plane in two different technology nodes. The targeted power supply voltage on the device plane is detected and set in less than 80 ns as shown through SPICE simulation.

In the second circuit implementation of the power module, frequency conversion through a clock divider is used to compare with an off-chip clock source operating at 25 MHz. A frequency comparator controls a voltage conditioning circuit that generates a precise reference voltage for an on-chip LDO or buck converter. Correct power supply voltage detection and clamping is demonstrated through circuit simulation for two device planes, one in 22 nm and the other in 45 nm. The power module is capable of setting the power supply voltage of a device plane ranging from 0.7 V to 2.5 V. The reference voltage is within 1% of the targeted power supply voltage, as indicated by simulated results.

Chapter 9: Evolving On-Chip Power Delivery through Particle Swarm Optimization

"Dumb parts, properly connected into a swarm, yield smart results."

– *Kevin Kelly, New Rules for the New Economy (1999), p.13.*

The advent of multi-core and many-core platforms with heterogeneous architectures has introduced new challenges for dynamic power management. A heterogeneous architecture offers higher energy efficiency for performance intensive data centers executing deep learning workloads. However, the power delivery through on-chip voltage regulators (OCVR) for heterogeneous many-cores is an ongoing and complex research topic [14, 95]. The optimal energy efficiency of a system is achieved with distributed and heterogeneous OCVRs that are interconnected through the power delivery network (PDN). A large and sustained current demand in GPUs subjects the OCVRs to increased aging and a higher susceptibility to process variation and noise. In addition, due to the finite latency of the OCVRs to react to changes in the load current, large dI/dt events in the GPU and cores results in power supply noise.

State of the art GPUs operate with off-chip voltage regulators [49], which leads to higher latency when reacting to changes in load current. In addition to high performance CPUs and GPUs, domain specific ASICs developed to accelerate deep learning applications such as the Google tensor processing unit (TPU) [12] also require advanced run time power management to mitigate timing errors due to process, voltage,

and temperature (PVT) variation and aging. The notion that deep neural networks (DNNs) are robust against PVT variations has recently been discredited [174]. DNNs with large systolic array multipliers such as the TPU, along with error resilient circuit techniques, improve the energy efficiency of the system through per layer voltage scaling [175].

The existing power management strategies for multi-core systems employ a central power management unit (PMU) that controls the operating voltage (and frequency for DVFS) of the cores and the core connectivity fabric. The decision to scale the voltage and frequency is executed by the operating system. The PMU provides the data gathered by the on-chip sensors (operating voltage, current, and/or temperature) to the operating system governor. In addition to the data provided by the physical sensors, the activity counters implemented in the architecture of the processor provide additional guidance on the appropriate operating voltage and frequency to the governor. A centralized power management strategy does not scale well as the number of processors in an IC increases. As shown in [176], the energy savings from the implementation of an active voltage guard-band in an IBM POWER7 diminishes as the number of active cores increases.

The design of the power distribution network is based on the optimization of a static voltage margin. A fixed voltage margin or guard-band is added to the power supply voltage to compensate for noise induced by the current drawn through the parasitic impedance of the PDN of the integrated circuit, package, and board in addition to the finite IR drop from the voltage regulator to the load circuits. In sub-20 nm nodes, the increase in process variation, the complex power-thermal interactions, and,

most importantly, the reduced voltage margin between the transistor operating voltage and threshold voltage requires novel methodologies for the correct design of the PDN.

In addition to technology scaling, FinFET based circuits result in higher current densities, which leads to an increase in dynamic power consumption, voltage noise, and thermal density [67]. Electromigration within the interconnect of the PDN is a growing concern for FinFET based circuits. Power management techniques such as power gating increase the inrush current, and therefore, the power supply noise. Optimization of decoupling capacitors alone does not effectively address the increased power supply noise [126].

Due to the complex interdependence of the various design challenges faced to properly and efficiently deliver power in sub-20 nm technology nodes, it is difficult to produce a robust and cost effective PDN using existing design techniques. Techniques based on vector-less peak power lead to an overcompensated and expensive PDN [177]. Vector or stimulus based PDN design with analysis of the peak power consumption provided through emulators is computationally expensive. Machine learning (ML) based techniques have recently been applied to the design of the PDN [178] to minimize the required routing resources while meeting the constraints for IR drop and electromigration. Another approach that utilizes ML algorithms for the design of the PDN leverages the power profile from a prior completed circuit to design the PDN of a current circuit. The two circuits may differ in functionality, but learning algorithms capture the physical characteristics of the circuits to effectively execute an informed decision that optimizes the PDN. With the rapid advancement in ML techniques,

the objective is to replace decisions in the design of an integrated circuit made based on human experience with ML algorithms. However, ML techniques applied to the PDN during the design phase can not mitigate the impact of circuit aging and power supply noise at run-time.

To effectively address the challenges faced in the delivery of power to circuits in advanced technology nodes, a run-time learning technique is needed for reliable and cost and energy efficient power delivery. In this chapter, an evolving on-chip voltage assignment is proposed and implemented with distributed OCVRs. The on-chip PDN self-learns and regulates the local voltages intelligently and autonomously to minimize the voltage guard-band without inducing any timing failures. Latched tap delay lines are integrated to sense the variation in the timing margin of critical paths, providing the collected data to the on-line learning algorithm. The data from the circuit-level sensor is directly provided to the system to reduce the execution latency of power modes including DVFS and increase the accuracy of the prediction of the optimal voltage and frequency for a given workload.

The main contributions of this work are:

- The development of a circuit technique for the adaptive voltage assignment to processing elements. The evolving voltage assignment is implemented with distributed on-chip voltage regulators of which the reference voltage is set through a particle swarm optimizer.
- The run time assignment of the power supply voltage compensates for the majority of second order effects limiting the reliable operation of processing cores

developed in sub-20 nm technologies with 3-D multi-gate transistors.

- This is the first work to apply machine learning at run time for the power management of processing elements that is contained completely in the circuit layer. Prior work on machine learning techniques for on-chip power delivery are applied during the design phase of a circuit [178]. Alternatively, machine learning techniques for power management and/or energy efficiency that execute partially at run-time rely on system or architectural level data for learning and inference [179].

The remainder of the chapter is organized as follows. The proposed particle swarm optimization based power management methodology is described in Section 9.1. Simulation results of an implementation of the runtime power management techniques are provided in Section 9.2. Summary of the work is provided in Section 9.3.

9.1 Learning algorithm for an evolvable PDN

Algorithms that optimize the control of the distributed on-chip voltage domains exist. Classical optimization methods including linear programming, non-linear programming, Newton's method, quadratic programming, and sequential unconstrained minimization assume that the variable being optimized is continuous, which yields local optimum solutions. The on-chip power delivery system with OCVRs contains both discrete and continuous control variables. Applying techniques for continuous variables to discrete variables results in both an increase of the objective function and in violations of inequality constraints. Evolutionary programming methods includ-

ing simulated annealing, genetic algorithm (GA), tabu search, and partical swarm optimization are better suited for discrete variables and non-differential objective functions [180].

Genetic algorithms are based on Darwinian theories of evolution and use processes analogous to genetic recombination and mutation to promote the evolution of a population that best satisfies a predefined objective. The selective crossover process involves choosing fit individuals to produce additional offspring, which improves the average result as the algorithm progresses. Subsequent mutations of the offsprings add diversity to the population and explore new areas of the search space of the parameter. Genetic algorithms have been extensively used in circuit design, particularly for high speed clock distribution [181] and post-silicon tuning of the clock delay [182]. A GA based method to determine the workloads that consume the peak power in a core has been implemented by ARM for the Cortex series of processors [183]. The primary disadvantage of applying a GA for run-time applications is an increase in the latency to optimize the circuit and architecture parameters as the search space increases. The increased latency and the globally sub-optimal results across the search space are limitations for other evolutionary techniques including simulated annealing and ant-colony optimization.

The particle swarm optimizer (PSO), however, offers a robust and simple implementation that produces superior results as compared to other evolutionary algorithms. Prior research has shown that the PSO offers different routes through the problem hyperspace as compared to the GA and other optimization algorithms [184, 185]. The low overhead to store results during each iteration of the algorithm

and the simplicity of the circuit implementation make the PSO algorithm an ideal choice for run-time control of the power supply voltages.

9.1.1 Voltage guard-band modulation based on particle swarm optimization

The particle swarm optimizer operates on a set of *particles* $p_1, p_2, \dots, p_n$, where the position of each particle x_i in a D dimensional hyperspace represents a potential solution to the optimization problem. For a given particle p_i , the position and velocity at time t are represented as, respectively, $x_i(t) = (x_{i,1}(t), x_{i,d}(t), \dots, x_{i,D}(t))$ and $v_i(t) = (v_{i,1}(t), v_{i,d}(t), \dots, v_{i,D}(t))$. The current best position for particle p_i is recorded as $P_{best} = (P_{i,1}(t), P_{i,d}(t), \dots, P_{i,D}(t))$. The best position among the entire particle population is recorded as G_{best} . The velocity and position of a particle constantly change based on both the experiences of the particle and the experiences of the other particles in the swarm. The position and velocity of the particles are updated as given by (9.1) and (9.2), respectively. In (9.1), w is the inertia weight, ϕ_1 and ϕ_2 are the learning factors, and ρ_1 and ρ_2 are random functions in the range of $[0,1]$.

The inertia weight w is applied to constrain the influence of past velocities on the current velocity of a particle. A large w enables a wider exploration of the hyperspace, whereas a small w results in a more local exploration to fine tune the current search results. A linear decrease in w with time (decreasing w strategy) yields near optimal results with the least iterations [186]. The cognitive parameter ϕ_1 provides a weight to the prior velocity of a particle when determining the current velocity. The social parameter ϕ_2 provides a weight to the swarm when determining the new velocity of

a particle. The values w , ϕ_1 , and ϕ_2 , therefore, set a procedure for the exploration of the hyperspace D .

$$\begin{aligned} v_{i,d}(t+1) = & w \times v_{i,d}(t) + \phi_1 \times \rho_1 \times (P_{i,d}(t) - x_{i,d}(t)) \\ & + \phi_2 \times \rho_2 \times (G_{i,d}(t) - x_{i,d}(t)) \end{aligned} \quad (9.1)$$

$$x_{i,d}(t+1) = x_{i,d}(t) + v_{i,d}(t+1) \quad (9.2)$$

9.1.2 Problem formulation with PSO

The current method to assign a supply voltage to a power domain consisting of various functional units is based on the addition of static voltage guard-bands. The appropriate guard-band is determined by assuming a pessimistic worst case analysis of the noise margins, which leads to the sub-optimal energy efficiency of the system. An implementation of the proposed methodology that applies the PSO to dynamically adjust the voltage guard-band at run-time is shown in Fig. 9.1. An SMT processor is shown with distributed on-chip voltage regulators supplying current to the various functional units. Distributed timing sensors measure the timing margin at the operating frequency of the circuit and transmit the data to the power management unit controlling the voltage reference of the OCVRs. The reference voltage of each OCVR is independently modulated based on the optimized values provided by the on-line PSO algorithm. Alternatively, to reduce the complexity of implementing the circuit that generates the reference voltage for each OCVR, the PSO provides a global

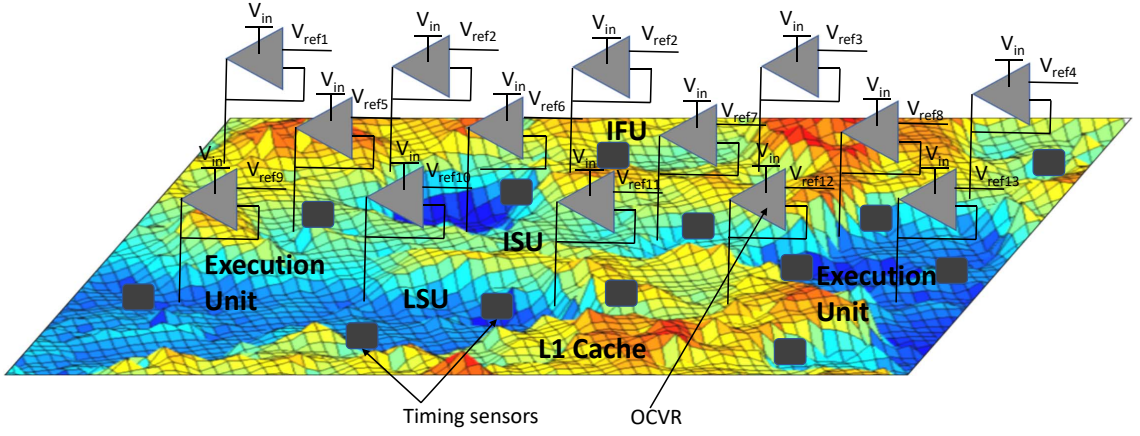


Figure 9.1: Distributed on-chip voltage regulators for a given voltage domain. The reference voltages to the distributed voltage regulators is obtained through the on-line particle swarm optimizer. The functional units of a typical SMT processor are shown as the load circuits in the voltage domain. The timing sensors are distributed across the domain to guide the PSO.

optimum voltage applied as a common reference to all OCVRs.

The optimization of the voltage guard-band for a given voltage domain with distributed on-chip voltage regulators is formulated as described by (9.3)-(9.6). The objective function is to maximize the energy efficiency of each voltage domain as given by (9.3):

$$\text{maximize } \eta_{energy} = f(\delta_1, \delta_2) \quad (9.3)$$

$$\text{s.t. } VID_{min} \leq \delta_1 \leq VID_{max} \quad (9.4)$$

$$V_{min} \leq \delta_2 \leq V_{max} \quad (9.5)$$

$$T_{margin} = g(\delta_2) \geq 0 \quad (9.6)$$

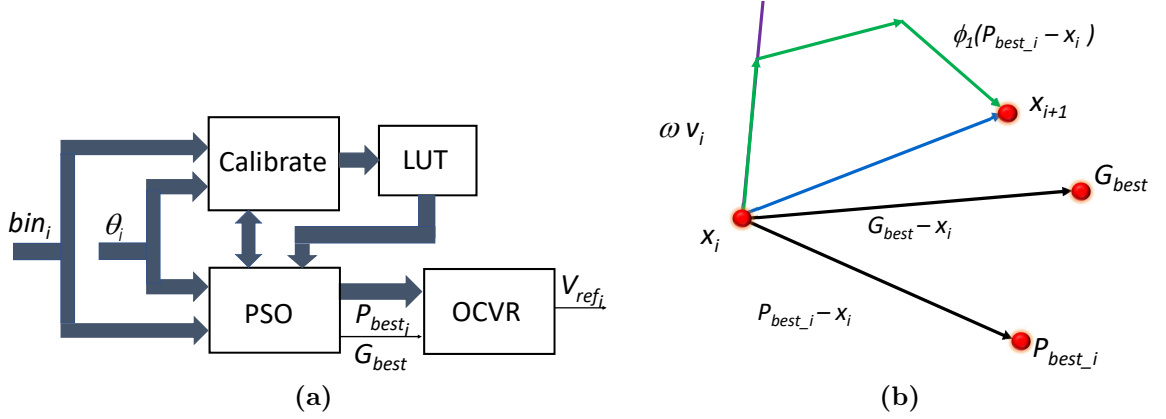


Figure 9.2: Assignment of a new voltage (position) to each OCVR (particle) shown through a (a) functional block diagram of the PSO algorithm and (b) a vector diagram.

The energy efficiency η_{energy} is a function of the control variable δ_1 and the dependent variable δ_2 . The control variable δ_1 represents the discrete reference voltage V_{ref_i} assignment to the OCVR p_i , while the dependent variable δ_2 represents the continuous output voltage of the OCVR. In a two tiered configuration of the on-chip power delivery network, the input voltage to the OCVRs is provided as a dependent variable to the optimization problem. The constraint given by (9.4) restricts the range of the reference voltage assignments to the programmable voltage identification (VID) levels of the power management circuit generating the voltage reference. The constraint given by (9.5) restricts the output voltage of the OCVRs in a given domain to the limit in the operating voltage of the MOS transistors set by the fabrication technology. The constraint given by (9.6) restricts the timing margin determined with the distributed timing sensors to a positive value. The timing margin of a critical path is inversely proportional to the set power supply voltage δ_2 of the OCVR.

9.1.3 Evolving voltage assignment

The distributed OCVRs operate as a swarm to locally optimize the operating voltage while applying the smallest needed guard-band to prevent timing violations on the local critical paths to compensate for aging related degradation in both the load and OCVR circuits. The optimization is completed through the computation of P_{best} and G_{best} , as given by, respectively, (9.7) and (9.8). The P_{best} of a particle (OCVR) is a function of the sensed time delay from the nearest timing sensor. The P_{best} values are the local optimal voltages accounting for local phenomenon including IR drop, dI/dt noise, process and aging induced V_{th} degradation, and hotspots. The G_{best} for the swarm of particles is the maximum P_{best} value obtained across all OCVRs. The P_{best} and G_{best} are functions of time, providing the lowest power supply voltage to the circuit without resulting in timing violations.

$$P_{best_i} = f(Temp(t), V_{noise}(t), V_{th_aging}(t), W_{load}(t)) \quad (9.7)$$

$$G_{best} = max(P_{best_i}) \quad (9.8)$$

The additional steps required to implement the PSO while designing a system with multiple processing cores are illustrated in Fig. 9.3(a). Existing statistical static timing analysis tools are applied to each voltage domain to determine the set of timing paths that exhibit high delay or are statistically likely to evolve as the paths with the maximum delay as the circuit ages. The set of likely timing paths are termed as potential critical paths (*PCPs*). A sensor is integrated within each *PCP* to monitor

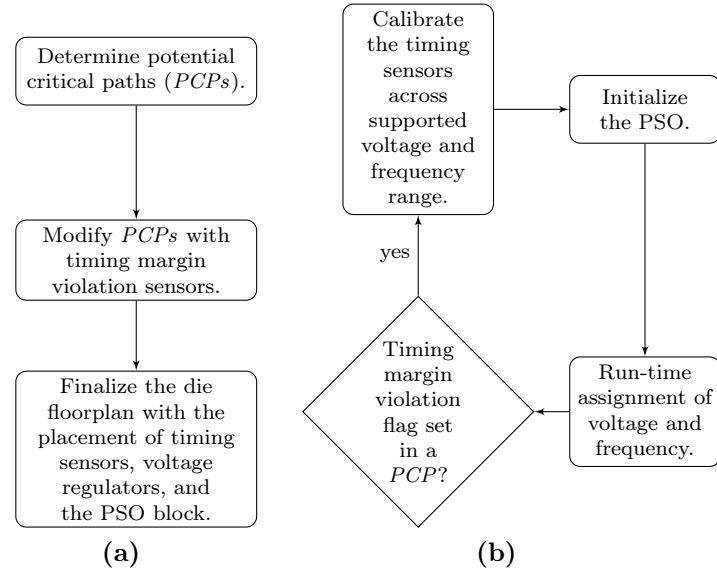


Figure 9.3: Methodology to implement the run-time adaptive voltage assignment using the PSO as described by (a) pre-silicon design steps, and (b) post-silicon process steps.

violations of the timing margin. The placement of the distributed voltage regulators and time based sensors is set with respect to the distribution of the *PCPs*.

The post-silicon procedure to execute the PSO is depicted in Fig. 9.3(b). At the first power up of the circuit, the distributed timing sensors are calibrated across the supported voltage and frequency ranges of the IC while executing a known workload that produces the lowest variation in the activity factor. The calibrated timing bins for each sensor at each frequency are stored on-chip in a look up table (LUT). Due to intra-die process variation, the calibrated bins vary for each timing sensor. The LUT is accessed by the on-line PSO, which compares the latest output from the timing sensor with the calibrated data in the LUT. As long as there is no violation in the timing margin (θ_i in Fig. 9.2(a)) of any of the PCP_i and the captured timing

bins are lower than the calibrated bins, the P_{best} position is updated. In case of a violation of the timing margin of any of the *PCPs*, a recalibration of the time based sensors is performed. The inputs and outputs to the on-chip PSO block are shown in Fig. 9.2(a). The algorithms for the evolving assignment of the power supply voltages, which includes the routines *CALIBRATE* and *PSO*, is described by pseudo code provided in Algorithm 9.1a.

The P_{best} and G_{best} values evolve with time as well as with changes in environmental conditions, operating temperature, and circuit aging. The procedure to assign an updated voltage (position) to each OCVR (particle) is shown in Fig. 9.2(b). The combined effect of aging and environmental conditions on the load circuits, voltage regulators, and the time based sensors is compensated by the computed voltages of the PSO, which are applied to the distributed regulators. With no timing violations, the assigned voltage(s) mitigate the formation of thermal hotspots. As the system ages based on the executed workloads and environmental conditions, a number of re-calibrations of the sensors are performed until there are no further possible modifications (reductions or increases) to the supply voltage at a given frequency.

9.1.4 Timing and aging sensor data to direct the PSO

On-chip sensors are needed to inform and direct the decision of the on-chip PSO for the assignment of voltages to the swarm of OCVRs. An optimal selection and placement of the sensors is required to most effectively characterize the operating voltage, temperature, and frequency of each voltage domain. In addition, for the proposed PSO, aging sensors are integrated to further characterize the state of the

circuit as a means to prevent timing violations due to aging in potential critical paths (PCPs). The construction and calibration of the timing and aging sensors is described as follows.

Timing sensor

A timing sensor such as a latched tapped delay line [187] provides an elegant and simple means to quantify the variation in the captured clock edges propagating through a chain of buffers. The variation in the captured clock edges is a function of the clock jitter, operating voltage, and temperature, which, as described in Chapter 3, are interrelated. Therefore, the combined effect must be characterized rather than individually quantifying the operating temperature, voltage, or load current with integrated physical sensors. Advanced circuit implementations of timing sensors are implemented in commercial microprocessors to characterize the available timing margin of critical paths (critical path monitors) [188, 189]. The commercial sensors can be integrated with the run-time PSO.

In this work, a latched tapped delay line is designed in a 7 nm FinFET PTM [66] process and is used as the timing sensor of the *PCPs* distributed across the IC to characterize and bin the location of the clock edge, with results provided to the on-line PSO. The schematic of the delay line based timing sensor is shown in Fig. 9.4(a). The local clock signal for the given voltage domain in which the delay line is placed is applied to the buffer chain. The buffer (or bin number) at which the clock edge is captured provides a measure of the local physical and electrical characteristics of the circuit in the vicinity of the delay line. An illustration of capturing a clock edge is

shown in Fig. 9.4(b) as the edge propagates through the delay line when an increase in the V_{DD} occurs. The clock edges propagate through more inverters when the delay per bin ($Delay_{perbin}$) decreases with an increase in supply voltage or temperature. An increase in V_{th} due to process variation and aging reduces the number of bins the clock edge propagates through.

The delay of a minimum sized inverter designed in a 7 nm FinFET PTM process at a nominal voltage of 0.7 V is 2.02 ps. Optimally sized buffers are implemented to reduce the size of the inverter chain and ensure that the fifth edge of the propagating clock at the highest supported operating frequency of 4 GHz is captured reliably across all PVT variation. The output of each buffer is latched into master/slave flip-flops. The two consecutive latches that capture the opposite logical output of the buffers indicates the location of the propagating clock edge (rising or falling) in the buffer chain. The location of the fifth clock edge ($bin_i[e_5]$) is considered as the input to the PSO algorithm as the sensitivity to V_{DD} increases the deeper the clock signal propagates into the buffer chain [187].

Aging sensor

A technique to predict circuit failure is developed in [190] based on monitoring the transition of the output signal of a critical path and the detection of any transitions within the set timing interval of the guard-band. A signal transition detected in the guard-band interval implies that for the given input to the combinational logic, the critical path has slowed due to circuit aging and is close to generating a timing fault. A monitoring circuit is embedded into the output latch of a critical path.

The block diagram of the monitoring circuit, consisting of the delay element and the stability checker, is shown in Fig. 9.5(a). The Output Latch stores the result of the stability checker. The schematic of the delay element and the stability checker is shown in Fig. 9.5(b). The delay element introduces a lag in the complement of the clock signal ($Clock'$), which is applied to the critical path. The delayed $Clock'$ signal is provided as input to the stability checker, which checks for any change in the output of the critical path during the guard-band interval as shown in Fig. 9.5(c). The global *Monitor* signal activates the delay element, and therefore, the aging sensor. The sensor detects transistor aging without applying any error correction or recovery techniques. For the proposed on-chip PSO based voltage assignment methodology, an aging sensor is placed in each of the potential critical paths ($PCPs$) determined during the statistical static timing analysis of a voltage domain. The output θ_i for the stability checker, from each of the PCP_i , is provided as an input to the PSO as shown in Fig. 9.2(a).

Calibration of the timing sensor

At the beginning of life of an IC, the distributed latched tapped delay lines are calibrated and the results are stored in a LUT. The calibration is performed at a nominal temperature of 25°C. A workload is executed on the processing elements of the voltage domain being calibrated that generates the least variation in the voltage of the power supply for a constant activity factor. The location of the timing sensors provides a characterization of the process variation in a given voltage domain as the delay per bin amongst the timing sensors varies with differences in V_{th} . The variation

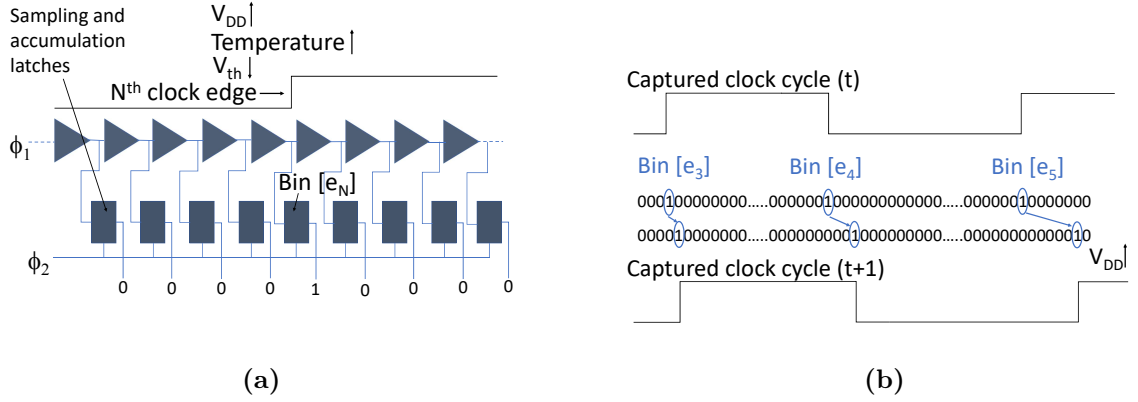


Figure 9.4: A latched tap delay line [187] in a 7 nm PTM FinFET technology [66] used as the timing sensor where (a) depicts the circuit schematic of the sensor and (b) the output of the latches capturing the shift in the clock edges due to an increase in V_{DD} .

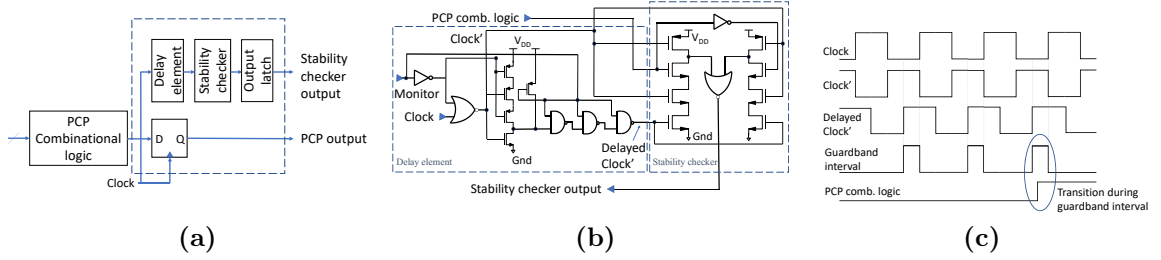


Figure 9.5: Aging sensor [190] built into each potential critical path (PCP) where (a) depicts a block diagram of the primary circuit components, (b) a circuit schematic of the sensor, and (c) a timing diagram depicting the detection of a violation due to aging.

in delay for a normal distribution (σ/μ of 0.9%) of V_{th} is shown in Fig. 9.6(c) for a 7 nm FinFET PTM technology [66] (distribution of V_{th} is shown in Fig. 9.6(d)). At each supported voltage level, the delay per bin ($Delay_{perbin}$) for each timing sensor i is calculated as the difference between the edges of one clock cycle ($bin_i[e_5] - bin_i[e_3]$), with the result stored in the LUT. The structure of the LUT with the corresponding stored results are listed in Table 9.1. The $bin_i[e_5]$ value corresponding to the nominal

voltage, as specified by the foundry for the given technology node is also stored in the LUT at the start of life for each timing sensor i as $calib_i$. The change in the $bin_i[e_5]$ location with operating voltage at a nominal temperature of 25°C and for a given V_{th} of 0.34 V is shown in Fig. 9.6(a). Due to TEI, the variation in the $bin_i[e_5]$ position with operating temperature for a supply voltage of 0.7 V and a V_{th} of 0.34 V is shown in Fig. 9.6(b). In both cases the $bin_i[e_5]$ location is given by the inverter at which the fifth clock edge is detected. The sensitivity of the latched tapped delay line to process, aging, voltage, and temperature, therefore, proves ideal as an on-chip sensor to direct the PSO algorithm.

During the operational lifetime of the IC, subsequent calibrations are performed if a violation in the timing margin is detected by the aging sensor. The calibration carried out during the lifetime of the IC is provided by the *CALIBRATE* procedure of the PSO as described in Algorithm 9.1a. The *Delayperbin* value D_1 corresponding to the current best voltage assignment P_{best_1} of an OCVR is compared with the updated calibration data obtained at each supported operating voltage. The updated voltage P_{best_2} corresponding to D_1 is assigned as the new reference voltage to the OCVR.

9.2 Simulated Results of PSO voltage assignment

The feasibility of run-time voltage assignment through the PSO is analyzed for SMs in a GPU and cores in a chip multi-processor (CMP). The run time PSO is, however, applicable to any circuit with on-chip distributed voltage regulators. The PSO algorithm, as given by Algorithm 9.1a, is implemented in MATLAB and in Verilog. The MATLAB model is used to characterize variations in the load current

Algorithm 9.1a Evolving power supply voltage assignment through particle swarm optimization.

Inputs:Set of aging sensors in each potential critical path (*PCP*): Θ Timing sensors output: $bin(n \times m)$ for n number of timing sensors and m timing sensor precisionSystem clock: clk **Outputs:**Local best voltage assignment: $P_{best}(n \times p)$ for n number of distributed voltage regulators and p bit VID codeGlobal best voltage assignment: G_{best} , p bit VID code

▷ Timing sensor calibration at beginning of life and at timing margin violation

procedure CALIBRATE(bin, clk, P_{best})**for** each $bin_i, i \in [1:n]$ **do****for** each $v_j \in VID_j$ **do**

▷ Compute the propagation delay per bin in the timing sensor

 $Delayperbin_j \leftarrow clk \times (bin_i[e_5] - bin_i[e_3])^{-1}$ **if** $calib_count == 0$ **then****if** $v_j == v_{nom}$ **then** $calib_i \leftarrow bin_i[e_5]$ $x_i \leftarrow v_j$ $calib_count++$ **else**▷ Search the LUT for the row with voltage equal to the P_{best_i} **if** $v_j == LUT_i(k, 1)(P_{best_i})$ **then** $D_1 = LUT_i(k, calib_count)(Delayperbin_k)$ **if** $Delayperbin_j == D_1$ **then** $calib_i \leftarrow bin_i[e_5]$ $x_i \leftarrow v_j$ $calib_count++$ ▷ Push the $Delayperbin$ computed per voltage in the LUT stack $LUT_i(j, calib_count) \leftarrow Delayperbin_j$ $P_{best_i} \leftarrow x_i$ $v_i \leftarrow x_i$ $G_{best} \leftarrow \max(P_{best_i})$ **return** $x, v, calib, P_{best}, G_{best}$ **procedure** PSO(Θ, bin, clk, x, v)▷ V_{max}, V_{min} : Technology imposed voltage limit for the transistor.

▷ Particle velocity limits

 $v_{max} = \gamma^*(V_{max} - V_{min});$ $v_{min} = -v_{max};$ ▷ $\gamma \in [0 \ 1]$ ▷ Constriction Coefficients: $\kappa \leftarrow 1, \phi_1 \in [0 \ 10], \phi_2 \in [0 \ 10],$ ▷ $\phi \leftarrow \phi_1 + \phi_2;$ ▷ $\chi \leftarrow \frac{2 \times \kappa}{|2 - \phi - \sqrt{\phi^2 - 4 \times \phi}|}$ $w \leftarrow \chi;$ ▷ Inertial Coefficient▷ $w_{damp} \in [0 \ 1];$ ▷ Damping ratio of inertia coefficient $c1 \leftarrow \chi \times \phi_1;$ ▷ Personal Acceleration Coefficient $c2 \leftarrow \chi \times \phi_2;$ ▷ Social Acceleration CoefficientAt every clk edge**if** $\theta_i \in \Theta == \text{TRUE}$ **then** ▷ timing violation on a *PCP*, recalibrate $(x, v, calib, P_{best}, G_{best}) \leftarrow \text{calibrate}(bin)$ **for** each $x_i, i \in [1:N]$ **do**

▷ Compute particle velocity

 $v_i \leftarrow w \times v_i + c1 \times \text{rnd} \times (P_{best_i} - x_i) + c2 \times \text{rnd} \times (G_{best} - x_i);$

▷ Apply velocity limits

 $v_i \leftarrow \min(v_i, v_{max});$ $v_i \leftarrow \max(v_i, v_{min});$

▷ Update particle position

 $x_i \leftarrow x_i + v_i;$ **if** $(bin_i[e_5] - calib_i > m'h0) \ \&\& \ (\theta_i \in \Theta == \text{FALSE})$ **then** $P_{best_i} \leftarrow x_i$ $w \leftarrow w \times w_{damp}$ $G_{best} \leftarrow \max(P_{best_i})$ **return** P_{best}, G_{best}

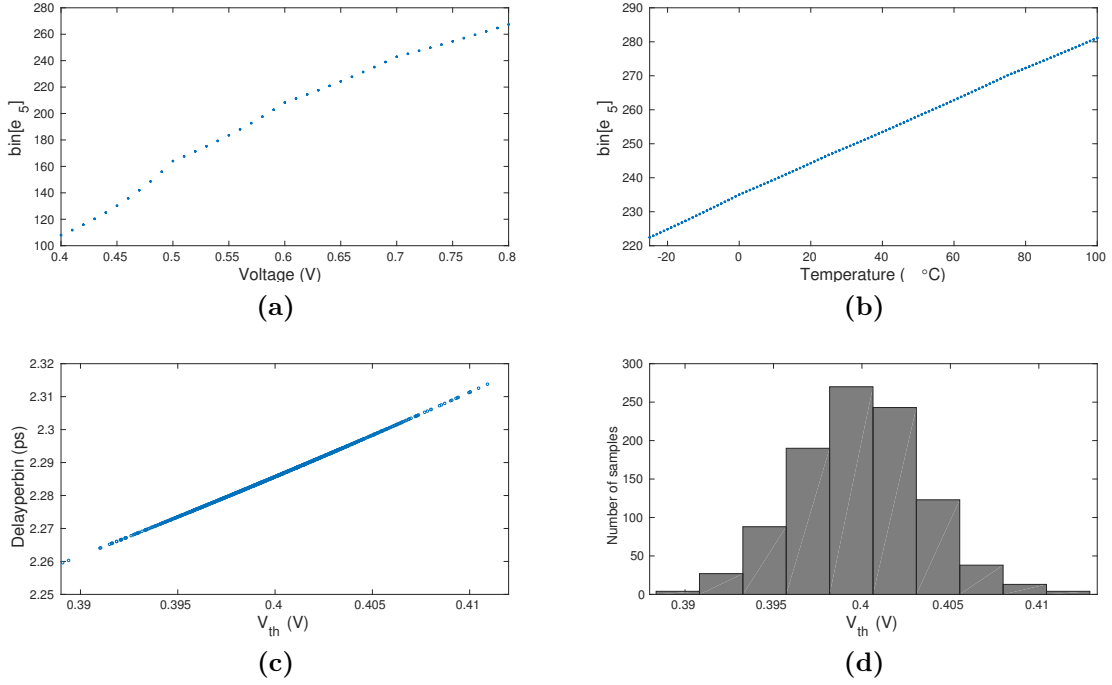


Figure 9.6: Response of the timing sensor at start of life with variation in (a) V_{DD} (at $T=25^{\circ}\text{C}$ and $V_{th}=0.34$ V), (b) operating temperature (at $V_{DD}=0.7$ V and $V_{th}=0.34$ V), and (c) intra-die V_{th} with a coefficient of variation σ/μ of 0.9 %. The distribution of V_{th} for a σ/μ of 0.9 % is provided in (d).

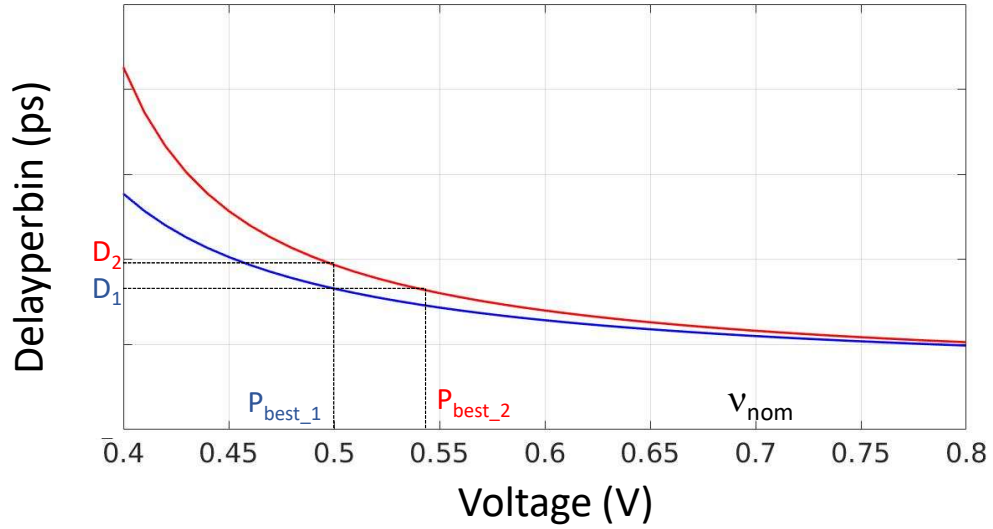
profile, power supply voltage, and threshold voltage. The verilog model is used in SPICE simulations along with Verilog-A models of the LDOs, timing sensors, aging sensors, and the on-chip power distribution network.

9.2.1 Simulation setup

A streaming multiprocessor (SM) of an NVIDIA GV100 GPU [49] is emulated with a constructed floor plan as shown in Fig. 9.8(a). Each SM in the GV100 is partitioned into four processing blocks, with each block containing 16 FP32 cores, 8 FP64 cores, 16 INT32 cores, a 64 KB register file, an L0 Icache, and two tensor cores. Three on-chip LDOs are considered per partition, which are roughly positioned within

Table 9.1: Results stored in a look up table (LUT_i) for the i^{th} timing sensor.

calib_count $\rightarrow$				
Voltage	$Delay_{perbin_i}$	$Delay_{perbin_i} \dots$	$Delay_{perbin_i}$	
V_{min}	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{best_1}	D_1	D_2	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{best_2}	$\vdots$	D_1	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
v_{nom}	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
V_{min}	$\vdots$	$\vdots$	$\vdots$	$\vdots$

**Figure 9.7:** An illustration of determining a new voltage each time the calibration procedure is invoked due to a flag raised by an aging sensor. The assigned new voltage results in the same delay per bin in the timing sensor as observed with the previous calibration.

the FP64/INT32 cores, the register file, and the tensor cores. All four partitions of the SM are considered to operate as a single voltage domain with 12 LDOs. The SM and the LDOs are designed in a 7 nm FinFET PTM process [66]. The power traces for the functional blocks of the SM and the SM floor plan are used as inputs to Voltspot 2.0 [191], which is a cycle accurate simulator. The voltage map across the power grid at discrete locations of the SM is obtained from Voltspot 2.0. The

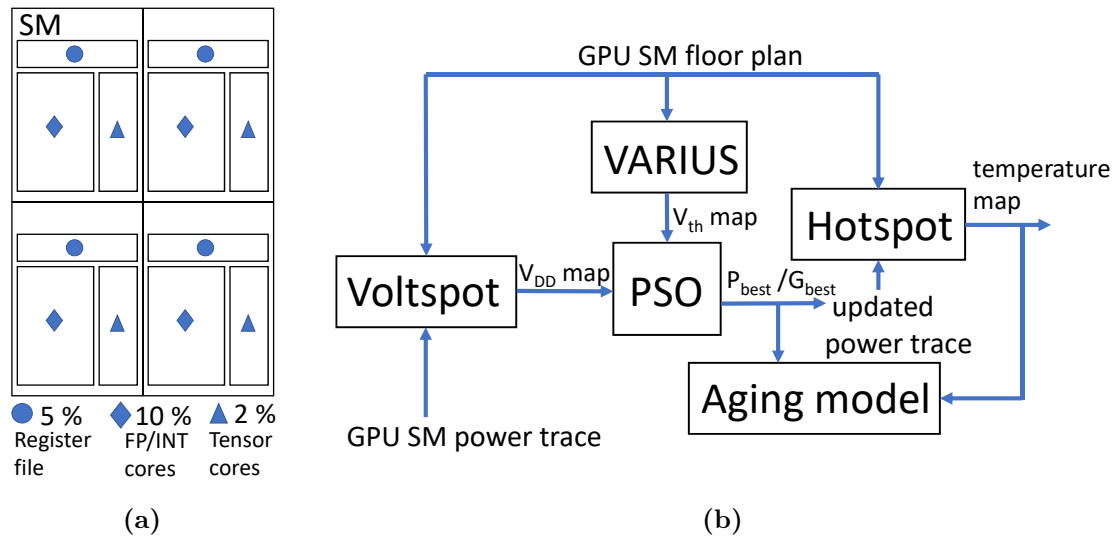


Figure 9.8: Simulation infrastructure to validate the PSO based voltage assignment algorithm where (a) depicts the floorplan of a SM used as an input to Voltspot and Hotspot and (b) the data flow diagram depicting the data parsed between the various tools used to characterize the PSO algorithm.

variation in V_{th} for the SM is determined using VARIUS [48], assuming a 0.9% σ/μ ratio and a spatial correlation parameter ϕ of 0.2. The voltage and the V_{th} maps are provided as inputs to the PSO model developed in MATLAB. The G_{best} obtained by the PSO per clock cycle is used to generate an updated power trace with the same activity factor as the original power trace inputted to Voltspot. The updated power trace is then applied as an input to the Hotspot [192] simulator to characterize the temperature across the SM. The simulation framework of the MATLAB based PSO is depicted in Fig. 9.8(b).

A four core chip multi-processor (CMP) which includes a verilog model of the PSO, is simulated in SPICE. The architectural parameters of the CMP are listed in Table 9.2. A set of 20 applications from the SPEC CPU2006 benchmark suite is used

Table 9.2: Architectural parameters of the core in a 16 core CMP.

Parameter	Value
Core clock frequency	4 GHz
Power supply voltage (V_{dd})	0.7 V
Issue, commit width	2
INT and FP instruction queue	16 entries
Load and Store Queue	16 entries
INT and FP Register File	48 entries
ROB size	48
L1 cache	32KB, 4-way
L2 cache	256KB

to determine the per cycle power dissipation [14] of the core in a 7 nm FinFET PTM process [66]. The power traces are used as piece wise linear waveforms to drive the current loads in the SPICE simulation. The current loads are connected to a circuit model of the (PDN), with each branch of the PDN designed as a series connection of a resistor and inductor. A total of 5 μ F of distributed on-chip capacitance is implemented. The 16 Verilog-A models of the LDO are distributed across the PDN such that there are four LDOs per core. A Verilog model of the latched tapped delay line is developed that accounts for the variation in delay with temperature, V_{DD} , and V_{th} , as shown in Fig. 9.6. A total of 16 delay lines are distributed in close proximity to the LDO models. The model of the circuit used for SPICE simulation is shown in Fig. 9.9.

9.2.2 Voltage assignment by the PSO algorithm

Simulations are conducted to characterize the execution of the PSO algorithm on a circuit with spatial and temporal variation in the power supply voltage and threshold voltage. Results from simulation of a voltage domain with 16 OCVRs are

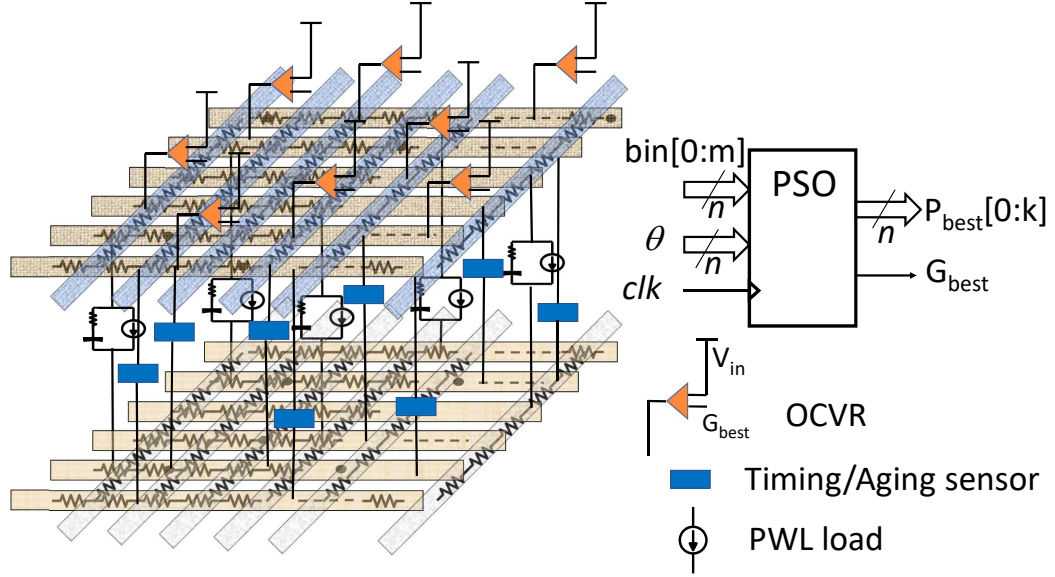


Figure 9.9: Model of the circuit used for SPICE simulation of a voltage domain that includes a circuit implementation of the PSO algorithm for voltage assignment of distributed OCVRs.

shown in Fig. 9.10. The voltage domain consists of 7 nm transistors that operate at a nominal voltage ν_{nom} of 0.7 V. The nominal $Delay_{perbin}$, simulated at $V_{DD} = \nu_{nom}$ and $T = 25^\circ \text{C}$ is 2.02 ps. A 20 ps timing margin is provided when implementing an aging sensor to monitor the timing of a critical path, where the initial delay of the paths at start of life is 200 ps. A σ/μ ratio of 0.9% is considered across the voltage domain to characterize process related variation. The clock frequency is set to 4 GHz. A uniform random power supply noise of 10% of V_{DD} is applied for a temperature is set to 25°C . The variation in the P_{best} values computed by the PSO for the 16 OCVRs for a time of execution of 1 μs is shown in Fig. 9.10(a). The evolving G_{best} value computed by the PSO for the given voltage domain is shown in Fig. 9.10(b).

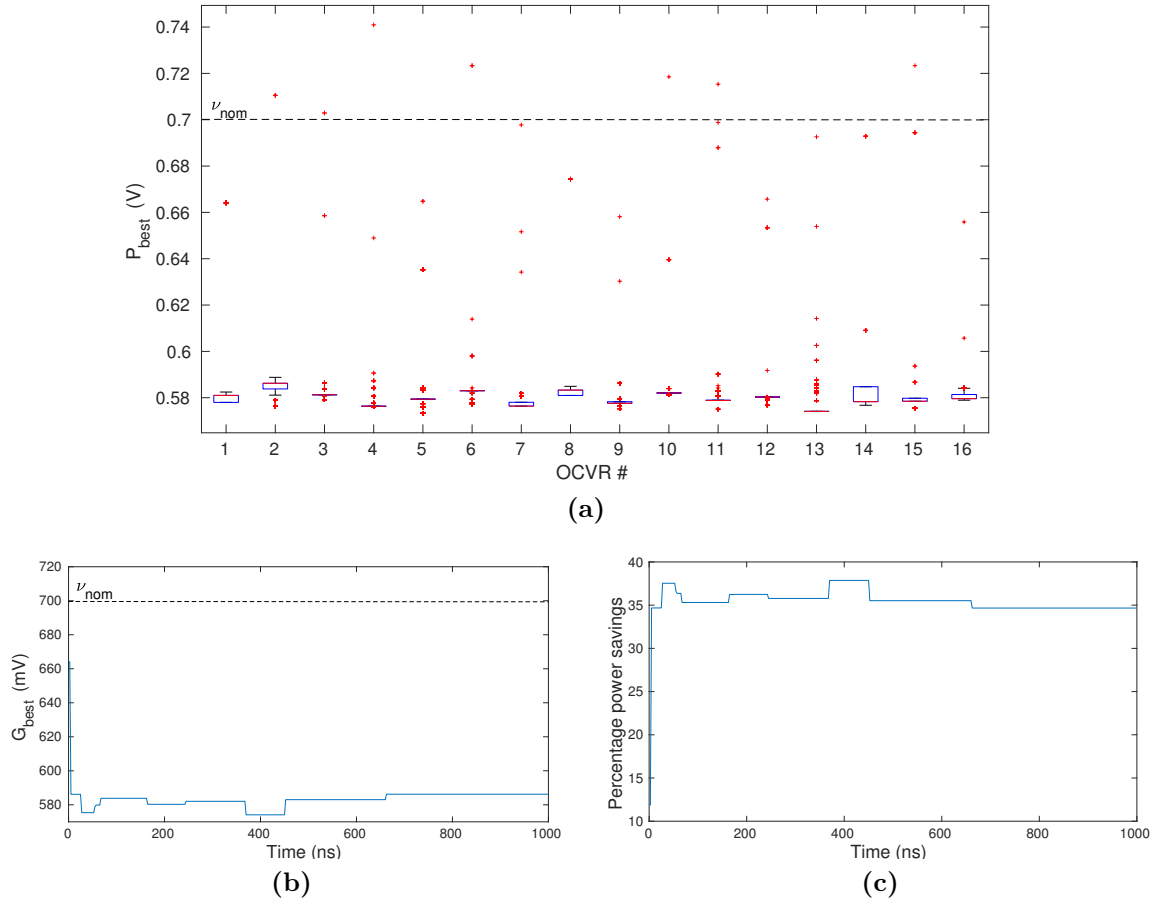


Figure 9.10: Results from the PSO based power supply voltage assignment algorithm for a voltage domain with 16 OCVRs. The voltage domain consists of gates from a 7 nm PTM process. The nominal voltage ν_{nom} for the transistors is 0.7 V. Included are the characterization of (a) the best voltage assignment per OCVR, (b) the global best voltage with time, and (c) the reduction in power consumption when applying the adaptive global voltage to the given voltage domain as compared to an assignment of ν_{nom} .

Even with voltage noise 5% greater than ν_{nom} , the PSO converges to a G_{best} value significantly lower than ν_{nom} without any timing violations in the critical path(s). The percentage reduction in the combined dynamic and static power consumption when applying the adaptive global G_{best} voltage assigned to a domain as compared to an assignment of ν_{nom} is shown in Fig. 9.10(c).

9.2.3 Power supply noise compensation

Prior research characterizing the power profile of a SM [193] determined that the caches are subject to the least amount of variation in power consumption. The FP and INT cores along with the register file (RF) are subject to large variations in power per cycle. The power variation characteristics of the NVIDIA tensor core are not publicly disclosed. However, if assuming a constant execution of matrix multiplication and addition operations by the tensor cores, the power consumption is assumed to be less variable as compared to the FP and INT cores.

Multi cycle power traces are generated for the FP/INT cores, the RF, and the tensor cores such that the induced voltage noise is, respectively, 10%, 5%, and 2% of the 0.7 V nominal V_{DD} . The OCVR supplying the RF also powers the L0 cache, warp scheduler, and the dispatch unit. Therefore, the combined variation in the activity of the partition is set to 5%. The activity pattern (temporal) and the placement of the circuit blocks (spatial) have a combined effect on the timing and magnitude of the power supply noise at any given location on the PDN [194]. The parameters of the on-line PSO are characterized to determine the effect on the exploration of the voltage search space with respect to the varying activity factors of the functional blocks of the SM. The personal acceleration coefficient (c_1) weighs the PSO decision more towards local noise events in the vicinity of the timing sensors, which are placed in close proximity to the OCVRs in the SPICE simulation. The reverse occurs for the social acceleration coefficient (c_2). The optimum voltage assignment is obtained when the personal and social acceleration coefficients are equal, which results in no

timing violations. The voltage assignment becomes more conservative (less variance between P_{best} values) as the c_2 coefficient is increased for a given c_1 . The results when setting both the personal (c_1) and social (c_2) coefficients equal to each other are shown in Fig. 9.11(a), whereas the results for the assignment of P_{best} when the PSO relies completely on the swarm (or social behavior) are shown in Fig. 9.11(b). The reduction in the variance of P_{best} when applying the voltage determined by the swarm optimization algorithm implies that blocks experiencing greater noise (overshoot due to LdI/dt) are assigned a lower voltage due to the influence of blocks with low to zero variation in activity. The reverse is true for the voltage assignment of blocks with low variation in activity, which provides less opportunity to reduce the voltage margin of such functional blocks.

In addition, if the operating system level workload scheduler provides data to the PSO from the architecture level activity counters, the inertial coefficient w is tuned based on the workload activity of the various functional blocks. A low value is assigned to w when more than one functional block exhibits high variation in activity factor, which ensures that the past voltage assignment of the PSO does not dominate the current assignment and the PSO searches for a solution that satisfies the transient power supply noise of the circuit.

9.2.4 Reduction in transistor aging

The aging induced degradation in V_{th} for an SM modeled as an NVIDIA GV100 is analyzed. An off-chip voltage regulator supplying current to the entire SM is considered as the baseline topology and is compared with the proposed technique imple-

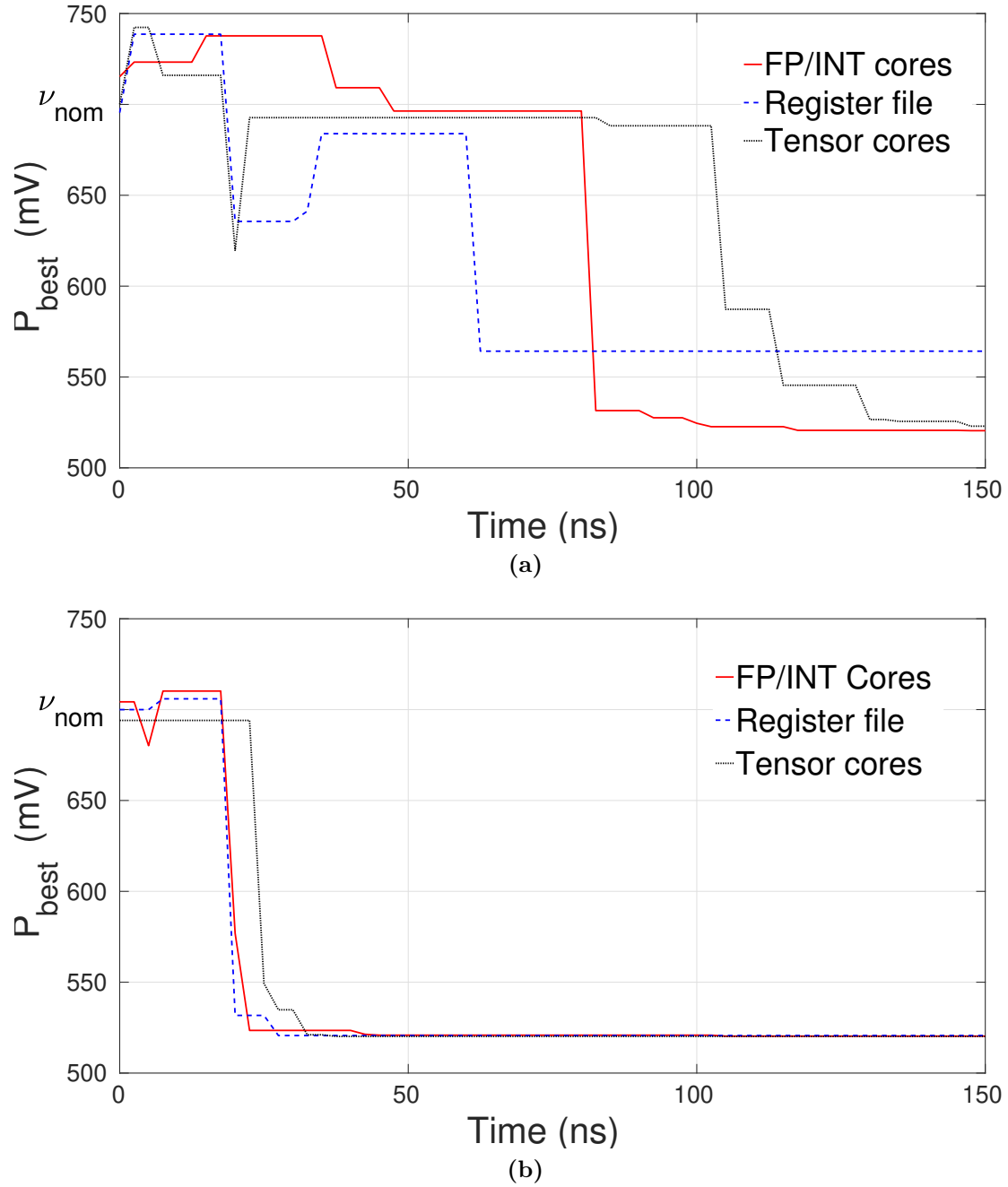


Figure 9.11: Characterization of the effect of the personal (c_1) and social acceleration (c_2) coefficients on the decision of the PSO (for varying circuit activity of the functional blocks). The evolving reference voltage assignment to a sub portion of a SM with three OCVRs is shown. The change in P_{best} applied to the OCVRs supplying the FP/INT cores, register file, and tensor cores is shown for decreasing levels of circuit activity for (a) $c_1 = c_2 = 2$, and (b) $c_1 = 0, c_2 = 2$.

menting a PSO based run-time voltage assignment of twelve distributed OCVRs, with placements as shown in Fig. 9.8(a). The selected aging model is validated through wafer measurements on a sub-20 nm FinFET-based ring oscillators [47]. The process and aging induced shift in the threshold voltage $\Delta V_{th}(t)$ is mathematically expressed as given by (9.9). The $\Delta V_{th}(t)$ has a normal distribution with a mean $\langle \Delta V_{thA}(t) \rangle$, which is the average shift in the threshold voltage attributed to BTI and is expressed by the power law given by (9.10). The technology and fitting parameters, A , κ , α , β , and γ in (9.10) are taken from prior work [59] and are listed in Table 9.3. The environmental and physical parameters required to calculate (9.10) include the temperature in Kelvin (K) θ , the total stress time in seconds t , the duty factor of the stress signal df , and the electric field across the gate oxide in V/m E_{OX} . The variance in the threshold voltage $\sigma_{\Delta V_{th}}^2(t)$ due to process variation and aging is given by (9.11), where $\sigma_{\Delta V_{th0}}^2$ is the variance due to process variation at the beginning of life of the SM. The SM is assumed to consist of an equal number of PMOS and NMOS transistors. The process induced variation in V_{th} is identical for the baseline and the PSO based SM.

$$\Delta V_{th}(t) = \mathcal{N}(\langle \Delta V_{thA}(t) \rangle, \sigma_{\Delta V_{th}}(t)) \quad (9.9)$$

$$\langle \Delta V_{thA}(t) \rangle \cong A e^{-\kappa/\theta} t^\alpha E_{OX}^\gamma df^\beta \quad (9.10)$$

$$\sigma_{\Delta V_{th}}^2(t) = \left(1 + \frac{\langle \Delta V_{thA}(t) \rangle}{0.1 \text{ V}}\right) \sigma_{\Delta V_{th0}}^2 \quad (9.11)$$

Table 9.3: Technology and power-law fitting parameters [59] used in the aging model.

Parameter	Description	NMOS	PMOS
γ	power-law exponent	5.2	3
α	power-law exponent	0.158	0.173
A	fitting constant	3.12e-2	2.02e-2
β	fitting constant	0.667	0.667
κ	fitting constant	50	50

The distributed OCVRs are implemented as LDOs. The aging of the PMOS header of the LDO is considered when determining the effect on the V_{th} due to the aging of the SM. As the output voltage from the LDOs is modulated by the PSO, variation occur in both the electric field across the gate oxide (E_{OX}) of the load circuits, which includes the distributed timing sensors, and the operating temperature. The updated temperatures across the SM are determined using Hotspot. The rate of aging for the baseline SM and the SM with PSO voltage assignment is calculated using the model given by (9.10) and (9.11) for the same stress time t and duty factor df . The variation in the V_{th} of an SM designed in a 7 nm FinFET process at the start of life and at the end of life (EOL) of 10 years for both the baseline SM and the SM with adaptive voltage assignment by the PSO are shown in Fig. 9.12. The analysis considers the G_{best} for the 12 OCVRs of a SM. Despite accounting for large power supply noise (10% on the FP/INT cores), the cumulative effect with time of the adaptive power supply voltage significantly reduces the rate of transistor aging, with a mean reduction in $\langle \Delta V_{thA}(t) \rangle$ of 40%. The improvement is due to the reduction in the applied electric field across the oxide E_{OX} of the transistors as compared to

the base line. The reduced temperature due to a lower applied voltage marginally improves the aging characteristics of the circuit. The reduced degradation in V_{th} due to aging is shown for both high- V_{th} and low- V_{th} 7 nm FinFET devices in Figs. 9.12(a) and 9.12(b), respectively.

9.2.5 Reduction in operating temperature

The thermal simulator HotSpot 6.0 [192] is used to characterize the effect of the evolving voltage assignment by the PSO on the temperature profile of the SM. Similar to Voltspot, the inputs to Hotspot are the SM architectural floorplan and the power trace. The floorplan of one GPU processing cluster (GPC) with 14 SMs [49] is considered for the temperature analysis. A series of differential equations are iteratively solved by Hotspot to compute the temperatures of the functional block. The average temperature of the area of the functional units is reported as an output. The transient temperatures for the RF, INT/FP cores, and tensor cores are determined through Hotspot by providing an updated power trace file computed using the G_{best} per clock cycle obtained by the PSO for 0.5 million clock cycles. Equal values of the personal (c_1) and social (c_2) acceleration coefficients are chosen for the analysis. The minimum, maximum, and mean reductions in temperature for each block in the floorplan as compared to the baseline are listed in Table 9.4. An average reduction of 5°C is observed across the RF, INT/FP cores, and tensor cores when a global best voltage assignment is applied to the circuit that is at least 20% less than the 0.7 V nominal voltage recommended for the 7 nm FinFET technology node. The self heating of the FinFETs is not captured through architectural level simulators including

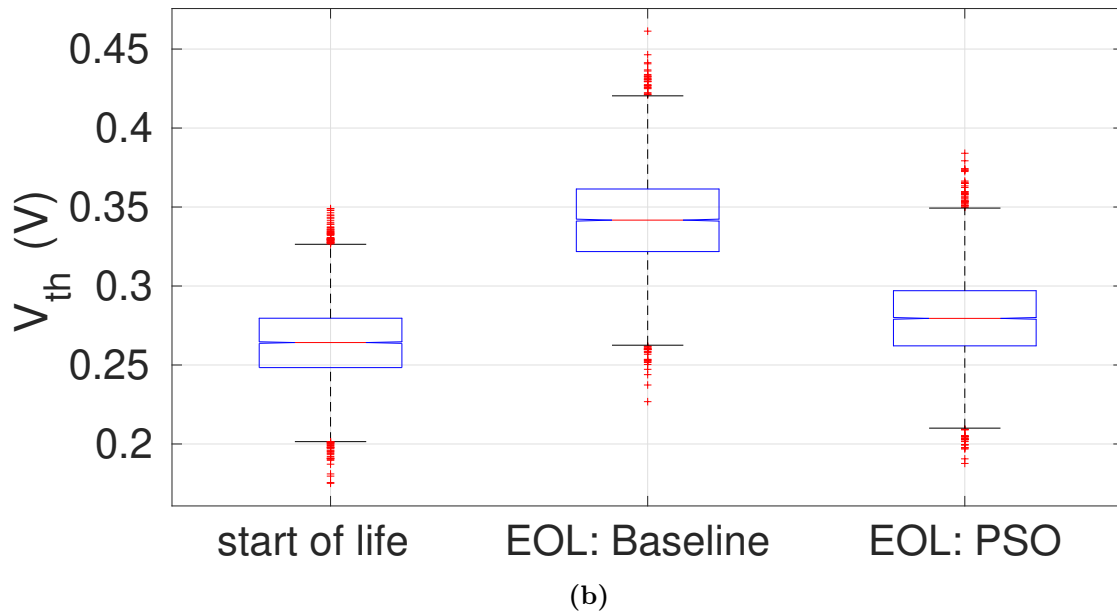
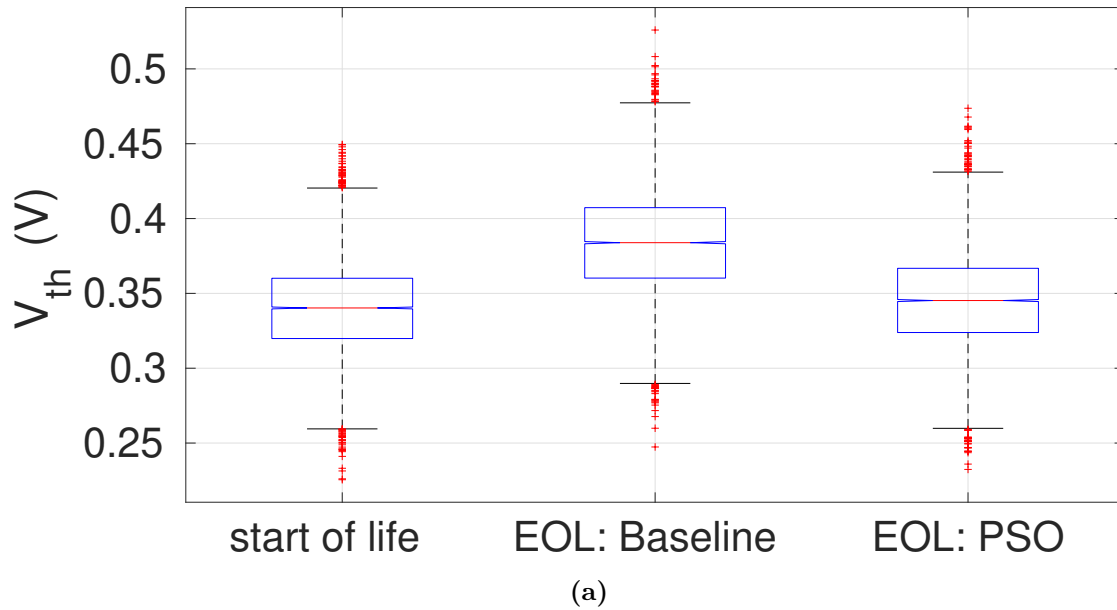


Figure 9.12: The reduction in aging induced degradation of V_{th} through adaptive voltage assignment to 12 OCVRs integrated in a SM modeled on the NVIDIA Volta GV100 [49] in a 7 nm FinFET process using (a) high- V_{th} and (b) low- V_{th} transistors. The end of life (EOL) in all cases is ten years.

Table 9.4: Analysis of the temperature of the functional blocks for the SM with voltage assignment through PSO as compared to a baseline implementing off-chip voltage regulators.

Functional unit	Average percentage reduction in the assigned voltage	Percentage reduction in temperature		
		Min	Max	Mean
RF	19.1	2.9	5.7	4.7
FP/INT cores	21.4	3.4	7.2	6.1
Tensor cores	22.1	2.3	5.1	4.2

Hotspot. As the self heating of the transistor and the local interconnects is directly proportional to the applied gate voltage, a reduction in the temperature due to self heating is expected when implementing the PSO based voltage assignment.

9.2.6 Hardware overhead of run-time PSO implementation

The overhead in area and compute cycles of implementing the on-line PSO for adaptive voltage assignment is estimated for the circuit shown in Fig 9.9, with the costs summarized in Table 9.5. The circuit and computational overheads to implement and execute the PSO are determined for a voltage domain with 16 distributed OCVRs supporting six distinct reference voltage levels, 16 delay line based timing sensors, and 12 *PCPs*. The computation time to determine a new P_{best} and G_{best} value for each OCVR is twenty clock cycles. The required size of the LUT increases with the operating age of the IC. As a means to reduce the size of the LUT, the allocated memory is reused after every two years of storing the calibration data. The fastest degradation in the threshold voltage of the OCVR and load circuit occurs in the first two years of the operating life of the IC. Beyond the first two years, the operating

Table 9.5: Circuit and computation overhead to implement the run-time PSO.

Parameter	Value
On-chip memory for LUT	512 B
Execution time of the <i>PSO</i> procedure in Algorithm 9.1a	20 clock cycles per OCVR
Aging sensors [190] in 12 <i>PCPs</i>	2K transistors
16 latched tapped delay lines [187] as timing sensors	6K transistors

temperature becomes a more critical parameter than the total stress time of the load circuit. Through the adaptive voltage assignment of the PSO, the reductions in voltage from the design time nominal value occur less frequently near the end of life of the IC. Therefore, the calibration results during the start of life of the IC do not need to be retained in the LUT.

9.3 Summary

An evolving on-chip voltage assignment to distributed on-chip voltage regulators is proposed. The reference voltage to each regulator is obtained through an evolutionary algorithm (particle swarm optimizer). The distributed voltage regulators serving a voltage domain work as a swarm and use the local data collected by the timing sensors to compensate for the effects of process variation, transistor aging in both the load circuit and the voltage regulators, and variations in the temperature across the die. The variation in the delay of the timing sensors is dependent on process, aging, temperature, and power supply noise, which provides optimal attributes to monitor the reliability and performance of the circuit. Through simulation on a 7 nm

FinFET technology, an average reduction of 35% and 5% in, respectively, the power consumption and operating temperature of the voltage domains was observed. In addition, the end of life of the circuit increases due to an average reduction of 40% in the aging induced variation in V_{th} .

The on-line PSO based power management technique is applicable to any computing platform including high performance general purpose processors, GP-GPUs, accelerators, or deep neural nets supporting on-chip voltage regulation. The efficacy of the PSO in suppressing power supply noise, reducing the operating temperature, and increasing the lifetime of the IC are demonstrated on a GPU SM and a multi-core CMP. The novelty of the method is in the modulation of the local supply voltages through on-chip and distributed voltage regulators to minimize the required voltage guard-band of a circuit while assuring no timing violations occur.

Chapter 10: Conclusions

A fundamental change in the field of computer engineering is underway due to the demise of Dennard's scaling and the slow down of Moore's law. Power and performance have been the primary drivers for research in the field of device technology, circuit design, computer architecture, and software and systems. The advent of many-core and heterogeneous computing has led to research addressing many of the challenges of power delivery to disparate cores while supporting power management techniques such as dynamic voltage and frequency scaling. As a consequence, on-chip voltage regulators (OCVRs) are inevitably needed in many-core systems. The circuit implementation of the voltage regulators including placement with respect to the cores and the connectivity with the cores is a complex optimization problem.

The focus of this dissertation is to optimize the power delivery network (PDN) for chip multi processors (CMPs) used for exascale computing to improve the energy efficiency of the system. There are many components to this research effort, not limited to the exploration of on-chip voltage regulator circuit topologies, power distribution networks in 2-D and 3-D ICs, run-time power management of multiple OCVRs in the presence of adaptive voltage, and thermal and power noise mitigation techniques. The concept of demand and supply side load management used in a *SMART grid* is applied to on-chip power delivery to infuse *intelligence* in the PDN composed of OCVRs, decoupling capacitors, power switches, and a power management unit.

OCVR topologies such as the linear drop-out (LDO) and step down buck converters are chosen for the analysis of the on-chip PDN that includes distributed OCVRs. The design of the PDN is optimized by accounting for the power conversion efficiency (PCE) of the OCVRs, which is, in the case of the buck converter, a function of the voltage inputted to and outputted from the OCVR, the load current drawn, and the switching frequency. The dissertation provides a very important conclusion for improving the energy efficiency of the CMP systems: Including the PCE in the optimization of the energy efficiency yields solutions with significant reduction in the energy consumption of the CMP. In Chapter 5, it is shown that with optimally sized OCVRs, the energy consumption of the CMP is reduced by up to 44% as compared to a CMP with OCVRs designed for the worst case. The optimal sizing of the OCVRs also reduces the on-chip foot print by at least 23%.

In addition to determining the optimal topology and output current rating of the OCVRs, a dynamic power delivery system, which reconfigures the connections between the OCVRs and the cores is developed. The run-time reconfiguration is controlled by an algorithm executed on the on-chip power management unit.

To further improve the energy efficiency of a CMP system with optimally sized distributed OCVRs provisioned to supply the average current demand of the circuit, a work load mapping heuristic is developed that maps the tasks to the cores such that the total current demand of the cores is always less than the total maximum output current rating of the OCVRs. A 100% scheduling of tasksets and assignment of DVFS levels for each core of a homogeneous CMP with task utilization factors of up to 0.55 is demonstrated. The workload mapping heuristic in conjunction with

the run-time reconfiguration of the power delivery network demonstrates cross-layer power management to ensure reliable and energy efficient operation of the CMP.

An analysis of power integrity in advanced technology nodes, specifically FinFET based processes, is described in Chapter 9. Challenges including circuit aging, process variation, power supply noise, and temperature hotspots are more pronounced in sub-nanometer nodes. In the literature, several techniques exist to mitigate such challenges, but the techniques address each sub-challenge in isolation. An evolving voltage delivery mechanism is developed that reduces circuit aging, power supply noise, and temperature hotspots, while accounting for process variation. The closed loop, run-time technique assigns voltages to the distributed OCVRs within the technology imposed guard-bands and without inducing any timing violation. A particle swarm optimizer (PSO) assigns the reference voltages to the OCVRs, which is provided data from distributed on-chip timing sensors that sense local variations due to power supply voltage, temperature, aging, and process. The simulation results for the PSO in 7 nm FinFET technology indicate an average reduction of 35%, 5% and 40% in, respectively, the power consumption, operating temperature, and threshold voltage drift for a voltage domain with a nominal power supply voltage of 0.7 V.

The fundamental cause of power supply noise in synchronous digital circuits is due to the switching activity of a circuit on clock edges. The thesis investigates the clock edge induced power supply noise, which cannot be completely mitigated with any of the existing power management techniques in literature. Existing techniques sense and react to clock edge induced power supply noise, which introduces a finite latency to implement the technique. Hyperabrupt junction varactors are demonstrated to

suppress clock edge induced power supply noise. The increase in capacitance as the voltage across the varactor drops is exploited to reduce the dependence of the voltage across the varactor terminals on the charge stored or released from the varactor. For the same amount of charge drawn, the voltage drop across series connected hyper-abrupt junction diodes is shown to be up to 60% less than an MIM or deep trench capacitor with the same capacitance.

Novel circuit techniques to detect and set the power supply voltages of disparate planes in a heterogeneous 3-D IC are also developed. Power supply voltage detection and clamping in the range of 0.7 V to 2.5 V is achieved for two device planes, one in a 22 nm and the other in a 45 nm process. The developed circuit techniques enable a “plug-and-play” approach to 3-D IC integration, as the disparate layers are manufactured in different fabrication facilities and integrated in a separate packaging facility.

Improved energy efficiency with increasing performance is possible in exascale data centers of the future through cross-layer power management. In addition, applying machine learning techniques such as the PSO to on-chip power management provides an innovative research direction towards improved energy efficiency for green computing. The on-chip power delivery methods in this dissertation improve the performance per watt advantage through novel circuit implementations and run-time circuits and systems techniques, which offer tremendous potential to achieve the 20 MW target set by the U.S. Department of Energy for exaflop computing.

Chapter 11: Future directions

The run-time techniques developed in this dissertation have been demonstrated through simulations on general purpose multi-core systems and GPUs. Through subtle modifications, the techniques can be applied to emerging processor architectures used in neural networks for artificial intelligence (AI) workloads as well as low power SoCs deployed as IoT sensors.

11.1 Energy efficient domain specific architectures

Processing elements such as GPUs yield results in a reasonable time when executing an AI algorithm, however, the demise of Moore’s law negates any further improvements in the performance per watt for such data intensive algorithms. Research in computer architecture is undergoing a renaissance period [195] with an expanding focus on domain specific architectures (DSA) to meet the increased computational demand of training deep neural networks. Alternately, there is an increased effort to push deep learning to the edge, which will strain the limits of battery operated devices.

Intelligence per joule is a new metric developed to characterize AI circuits. Current research objectives include the optimization of DSA architectures and domain specific programming languages [195] to improve the energy efficiency of the circuit. An unexplored research area with the potential to improve the cost-energy-performance of DSAs is power delivery through distributed on-chip voltage regulators. The run-

time learning for optimal on-chip voltage assignment and the reconfigurable power distribution network described in this dissertation provides a foundation for several research directions for energy management in DSAs. A few such plausible directions include:

1. Aggressive voltage under-scaling of high-performance DNN accelerators such as the Google TPU, without significant reduction in classification accuracy, has been proposed in literature [175]. The Google TPU includes a systolic array based architecture, which is structured as a 2-D grid of multiplication and accumulate (MAC) units. As the communication between the MACs is limited to only the nearest neighbor, the systolic array architecture reduces the routing and memory overhead as well as the latency due to communication between separate MACs. However, the entire systolic array operates as a single voltage domain. The reconfigurable PDN developed in Chapter 5 can be implemented for each layer of systolic array architectures. With the PSO based voltage delivery technique described in Chapter 9, per layer voltage under-scaling is feasible. As shown in [175], per layer voltage under-scaling achieves a 34% to 57% reduction in energy consumption on speech and image recognition benchmarks with less than 1% loss in classification accuracy.
2. Currently, training convolutional neural networks (CNNs) in a reasonable amount of time is done through a cluster of GPUs. Training is communication intensive and requires a fast and efficient fabric to transmit data between GPUs. The power delivery to the communication fabric of a CNN is a prime candidate for

application of the developed PSO based evolvable voltage assignment methodology. A technique to supply optimal voltages to the communication channels such that the energy efficiency of the DNN increases without degrading the performance is feasible using the developed on-line PSO based power management technique.

11.2 Power management of swarm of IoT sensors for military applications

Real time connectivity between command-center and armed force personnel is a fundamental requirement for any military operation. A high degree of situational awareness in combat zones along with robust and secure communication drive the need for advanced human to machine and machine to machine (M2M) communication. The growth in IoT electronics, which span multiple vertical markets, is driven by the need for M2M devices in advanced combat systems. The possible military applications for M2M devices range from equipment monitoring for early detection and diagnosis of faults to biological, chemical, and physical sensors. The deployment of devices in remote and hostile environments introduces unique challenges in battery life, ultra-low power operation, and the need for highly robust power delivery systems.

The long operational lifetime of electronics in the battle field requires a large mean time between failure and environmental robustness for circuits operating under adverse conditions. The integrated circuits found in these devices must, therefore, offer ultra-high energy efficiency, environmental sensing for long durations, and information processing. The two primary sources of power for M2M devices are batteries and/or

energy harvesting circuits. The DC-DC conditioning circuit converts and assures the correct DC voltage levels to each power domain within a device. The different circuit blocks that can include environmental sensors (biological, chemical or physical), external communication (RF subsystem and antennae), and information processing (heterogeneous cores) require multiple power supplies to deliver sufficient and stable current at varying voltage levels. Therefore, on-chip voltage regulators are needed to support the multiple analog and digital voltage domains in such heterogeneous IoT devices.

The extreme environmental conditions result in power and signal integrity challenges for IoT nodes. The run-time voltage detect and set techniques developed in Chapter 8 provide foundational work to develop a run-time power supply voltage drift detection and correction mechanism that counteracts the effect of PVT and aging induced voltage variation in M2M IoT nodes deployed in extreme environments.

The instantaneous power available from an energy harvesting system is variable. To minimize the energy consumption from the battery, the peak load current supported by the OCVRs must be lowered such that the ratio of the power harnessed from the energy harvester to the power sourced from the battery increases. The load circuits (processors, RF transceiver, and sensors) driven by an energy harvesting system, therefore, must consume energy at an appropriate rate that is efficiently supported by the implemented energy harvester. The run time OCVR clustering algorithm developed in Chapter 5 is applicable to IoT nodes that include energy harvesting systems for power generation.

11.3 Applications to hardware security

Non-invasive side channel attacks on digital circuits are potent tools to extract information as well as encryption keys in cryptographic blocks of an IC. Differential power analysis (DPA) and leakage power analysis (LPA) are two such non-invasive side channel attacks that are of growing concern in the field of hardware security [196]. DPA and LPA attacks utilize the correlation between, respectively, the input data and dynamic power and the leakage power consumption of the digital block under attack.

The run-time reconfigurable PDN developed in Chapter 5 provides a means to mask the load current profile from DPA and LPA attacks. The run-time reconfiguration of the connections between the load circuits and the distributed OCVRs reduces the correlation between the dynamic and leakage power dissipated by the load circuits from the power and ground pins supplying current to the global on-chip PDN to which the OCVRs are connected. The OCVR clustering and declustering algorithm can be suitably modified to counteract a DPA or LPA attack on the power grid controlled by the on-chip power management unit.

Bibliography

- [1] S. Ashby, P. Beckman, J. Chen, P. Colella, B. Collins, D. Crawford, J. Dongarra, D. Kothe, R. Lusk, P. Messina, and et al., “The Opportunities and challenges of exascale computing,” *Summary Report of the Advanced Scientific Computing Advisory Committee Subcommittee*, pp. 1–77, 2010.
- [2] R. Lucas, “Top ten exascale research challenges,” *DOE ASCAC Subcommittee Report*, February 2014.
- [3] W. Harrod, “A journey to exascale computing,” *High performance computing, networking, storage and analysis*, pp. 1702–1730, 2012.
- [4] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, “Design of ion-implanted MOSFET’s with very small physical dimensions,” *IEEE Journal of Solid-State Circuits*, Vol. 9, No. 5, pp. 256–268, October 1974.
- [5] B. Davari, R. H. Dennard, and G. G. Shahidi, “CMOS scaling for high performance and low power-the next ten years,” *Proceedings of the IEEE*, Vol. 83, No. 4, pp. 595–606, April 1995.
- [6] C. A. Mack, “Fifty years of Moore’s law,” *IEEE Transactions on semiconductor manufacturing*, Vol. 24, No. 2, pp. 202–207, January 2011.
- [7] K. Rupp, M. Horovitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten, “Years of microprocessor trend data,” <https://www.karlrupp.net/2015/06/40-years-of-microprocessor-trend-data>.
- [8] Advanced Micro Devices Inc., “AMD threadripper processors,” <https://www.amd.com/en/products/ryzen-threadripper>.
- [9] Intel Inc., “Intel core i9 desktop processor,” <https://www.intel.com/content/dam/support/us/en/documents/processors/intel-core-i9-comparison-chart.pdf>.
- [10] S. Fiore, M. Bakhouya, and W. W. Smari, “On the road to exascale: Advances in high performance computing and simulations—an overview and editorial,” 2018.
- [11] K. O’Brien, L. Di Tucci, G. Durelli, and M. Blott, “Towards exascale computing with heterogeneous architectures,” *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition*, pp. 398–403, March 2017.

- [12] N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, *et al.*, “In-datacenter performance analysis of a tensor processing unit,” *Proceedings of the International Symposium on Computer Architecture*, pp. 1–12, June 2017.
- [13] M. Davies, N. Srinivasa, T. Lin, G. Chinya, P. Joshi, A. Lines, A. Wild, and H. Wang, “Loihi: A neuromorphic manycore processor with on-chip learning,” *Proceedings of the IEEE Micro*, Vol. 38, January 2018.
- [14] D. Pathak, H. Homayoun, and I. Savidis, “SMART GRID ON CHIP: Work load balanced on-chip power delivery,” *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 25, No. 9, pp. 2538–2551, September 2017.
- [15] D. Pathak, M. Hajkazemi, M. Tavana, H. Homayoun, and I. Savidis, “Energy Efficient On-Chip Power Delivery with Run-Time Voltage Regulator Clustering,” *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1210–1213, May 2016.
- [16] I. H. Jiang, G. J. Nam, H. Y. Chang, S. R. Nassif, and J. Hayes, “Smart Grid load balancing techniques via simultaneous switch/tie-line/wire configurations,” *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 382–388, November 2014.
- [17] D. Pathak, M. Hajkazemi, M. Tavana, H. Homayoun, and I. Savidis, “Load Balanced On-Chip Power Delivery for Average Current Demand,” *Proceedings of the Great Lakes Symposium on VLSI*, pp. 439–444, May 2016.
- [18] W. Kim, M. S. Gupta, G. Wei, and D. Brooks, “System level analysis of fast, per-core DVFS using on-chip switching regulators,” *Proceedings of the IEEE International Symposium on High Performance Computer Architecture*, pp. 123–134, February 2008.
- [19] A. Chaehoi, M. Begbie, D. Weiland, and S. Scherner, “Multiple-output MEMS DC/DC converter,” *Proceedings of the IEEE Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS*, pp. 253–257, March 2011.
- [20] S. Ghandour, G. Despesse, and S. Basrour, “Design of a new MEMS DC/DC voltage step-down converter,” *IEEE International NEWCAS Conference*, pp. 105–108, June 2010.
- [21] M.D. Seeman, *A Design Methodology for Switched-Capacitor DC-DC Converters*, Ph.D. Thesis, University of California Berkeley, 2009.
- [22] J. Gjanci, *On-chip voltage regulation for power management in system-on-chip*, Master’s thesis, University of Illinois at Chicago, 2008.
- [23] I. Vaisband and E. G. Friedman, “Heterogeneous methodology for energy efficient distribution of on-chip power supplies,” *IEEE Transactions on Power Electronics*, Vol. 28, No. 9, pp. 4267–4280, September 2013.

- [24] E. Salman and E. G. Friedman, *High performance integrated circuit design*, McGraw Hill, 2012.
- [25] B. Allard, *Power systems-on-chip: practical aspects of design*, John Wiley & Sons, 2016.
- [26] S. Kose, S. Tam, S. Pinzon, B. McDermott, and E. G. Friedman, “Active filter-based hybrid on-chip DC–DC converter for point-of-load voltage regulation,” *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 21, No. 4, pp. 680–691, April 2013.
- [27] V. Kursun, S. G. Narendra, V. K. De, and E. G. Friedman, “Analysis of buck converters for on-chip integration with a dual supply voltage microprocessor,” *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 11, No. 3, pp. 514–522, June 2003.
- [28] M. Al-Shyoukh, H. Lee, and R. Perez, “A transient enhanced low-quiescent current low-dropout regulator with buffer impedance attenuation,” *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 8, pp. 1732–1742, August 2007.
- [29] H. Le, M. Seeman, S. R. Sanders, V. Sathe, S. Naffziger, and E. Alon, “A 32 nm fully integrated reconfigurable switched-capacitor DC-DC converter delivering 0.55 W/mm² at 81% efficiency,” *IEEE Journal of Solid-State Circuits*, pp. 210–211, February 2010.
- [30] E. J. Fluhr, S. Baumgartner, D. Boerstler, J. F. Bulzacchelli, T. Diemoz, D. Dreps, G. English, J. Friedrich, A. Gattiker, T. Gloekler, *et al.*, “The 12-Core POWER8 Processor With 7.6 Tb/s IO Bandwidth, Integrated Voltage Regulation, and Resonant Clocking,” *IEEE Journal of Solid-State Circuits*, Vol. 50, No. 1, pp. 10–23, January 2015.
- [31] 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,” *Proceedings of the IEEE Applied Power Electronics Conference and Exposition*, pp. 432–439, March 2014.
- [32] U. Zillmann, A. Schaefer, R. Saraswat, T. Kamgaing, P. B. Fischer, and G. Droege, “Integrated voltage regulators with magnetically enhanced inductors,” March 2018, US Patent 9,921,640.
- [33] M. Erturk, F. Mahmoudi, J. T. Doyle, R. V. Shenoy, and J. Kim, “Integrated passive on a substrate,” March 2017, US Patent App. 14/859,320.
- [34] N. Rahman, A. Parayandeh, K. Wang, and A. Prodic, “Multimode digital SMPS controller IC for low-power management,” *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 4–8, May 2006.

- [35] S. B. MahdaviKhah-Mehrabad, *Hybrid sc-inductive based topologies for low-volume high-efficiency switched-mode power supplies*, Ph.D. Thesis, University of Toronto, 2014.
- [36] S. M. Ahsanuzzaman, A. Parayandeh, A. Prodić, and D. A. Johns, “A building block IC for designing emerging hybrid and multilevel converters,” *IEEE Journal of Emerging and Selected Topics in Power Electronics*, Vol. 6, No. 2, pp. 500–514, November 2018.
- [37] J.F. Bulzacchelli, Z. Toprak-Deniz, T.M. Rasmus, J.A. Iadanza, W.L. Bucossi, S. Kim, R. Blanco, C.E. Cox, M. Chhabra, C. D LeBlanc, *et al.*, “Dual-Loop System of Distributed Microregulators With High DC Accuracy, Load Response Time Below 500 ps, and 85-mV Dropout Voltage,” *IEEE Journal of Solid-State Circuits*, Vol. 47, No. 4, pp. 863–874, April 2012.
- [38] C. Gonzalez, M. Floyd, E. Fluhr, P. Restle, D. Dreps, M. Sperling, R. Rao, D. Hogenmiller, C. Vezirtis, P. Chuang, *et al.*, “The 24-core POWER9 processor with adaptive clocking, 25-Gb/s accelerator links, and 16-Gb/s PCIe Gen4,” *IEEE Journal of Solid-State Circuits*, Vol. 53, No. 1, pp. 91–101, January 2018.
- [39] J. E. Lilienfeld, “Method and apparatus for controlling electric currents,” January 1930, US Patent 1,745,175.
- [40] S. Thompson, “MOS scaling: Transistor challenges for the 21st century,” *Intel Technology Journal*, 1998.
- [41] D. Hisamoto, W. C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T. J. King, J. Bokor, and C. Hu, “FinFET—a self-aligned double-gate MOSFET scalable to 20 nm,” *IEEE Transactions on Electron Devices*, Vol. 47, No. 12, pp. 2320–2325, December 2000.
- [42] Y. K. Choi, N. Lindert, P. Xuan, S. Tang, D. Ha, E. Anderson, T. J. King, J. Bokor, and C. Hu, “Sub-20 nm CMOS FinFET technologies,” pp. 19.1.1–19.1.4, December 2001.
- [43] G. K. Celler and S. Cristoloveanu, “Frontiers of silicon-on-insulator,” *Journal of Applied Physics*, Vol. 93, No. 9, pp. 4955–4978, May 2003.
- [44] R. Shen, S. Tan, and H. Yu, *Statistical performance analysis and modeling techniques for nanometer VLSI designs*, Springer Science & Business Media, 2012.
- [45] R. Thomas, K. Barber, N. Sedaghati, L. Zhou, and R. Teodorescu, “Core tunneling: Variation-aware voltage noise mitigation in GPUs,” *Proceedings of the Symposium on High Performance Computer Architecture*, pp. 151–162, March 2016.

- [46] T. Baji, “Evolution of the GPU device widely used in AI and massive parallel processing,” *Proceedings of the IEEE Electron Devices Technology and Manufacturing Conference*, pp. 7–9, March 2018.
- [47] P. Weckx, B. Kaczer, M. Toledano-Luque, P. Raghavan, J. Franco, P. J. Roussel, G. Groeseneken, and F. Catthoor, “Implications of BTI-induced time-dependent statistics on yield estimation of digital circuits,” *IEEE Transactions on Electron Devices*, Vol. 61, No. 3, pp. 666–673, January 2014.
- [48] S. R. Sarangi, B. Greskamp, R. Teodorescu, J. Nakano, A. Tiwari, and J. Torrellas, “VARIUS: A model of process variation and resulting timing errors for microarchitects,” *IEEE Transactions on Semiconductor Manufacturing*, Vol. 21, No. 1, pp. 3–13, February 2008.
- [49] NVIDIA, “NVIDIA Tesla V100 GPU Architecture,” <https://images.nvidia.com/content/volta-architecture/pdf/volta-architecture-whitepaper.pdf>, pp. 1–58, August 2017.
- [50] G. C. K. Y. Chen, K. Y. Chuah, M. F. Li, D. S. H. Chan, C. H. Ang, J. Z. Zheng, Y. Jin, and D. L. Kwong, “Dynamic NBTI of PMOS transistors and its impact on device lifetime,” *Proceedings of the IEEE International Reliability Physics Symposium*, pp. 196–202, March 2003.
- [51] D. Bergstrom, M. Hattendorf, J. Hicks, J. Jopling, J. Maiz, S. Pae, C. Prasad, and J. Wiedemer, “Intel’s 45 nm CMOS technology,” *Intel Technology Journal*, Vol. 12, No. 2, pp. 131–144, June 2008.
- [52] D. K. Schroder and J. A. Babcock, “Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing,” *Journal of Applied Physics*, Vol. 94, No. 1, pp. 1–18, June 2003.
- [53] S. Mahapatra, N. Goel, S. Desai, S. Gupta, B. Jose, S. Mukhopadhyay, K. Joshi, A. Jain, A. E. Islam, and M. A. Alam, “A comparative study of different physics-based NBTI models,” *IEEE Transactions on Electron Devices*, Vol. 60, No. 3, pp. 901–916, January 2013.
- [54] K. B. Sutaria, J. B. Velamala, A. Ramkumar, and Y. Cao, *Compact modeling of BTI for circuit reliability analysis*, Springer, 2015.
- [55] E. Mintarno, J. Skaf, R. Zheng, J. Velamala, Y. Cao, S. Boyd, R. W. Dutton, and S. Mitra, “Optimized self-tuning for circuit aging,” *Proceedings of the Conference on Design, Automation and Test in Europe*, pp. 586–591, March 2010.
- [56] Y. Yu, J. Liang, Z. Yang, and X. Peng, “NBTI and power reduction using a workload-aware supply voltage assignment approach,” *Journal of Electronic Testing*, Vol. 24, No. 1, pp. 27–41, February 2018.

- [57] S. Gupta and S. S. Sapatnekar, “GNOMO: Greater-than-NOMinal Vdd operation for BTI mitigation,” *Proceedings of the Asia and South Pacific Design Automation Conference*, pp. 271–276, January 2012.
- [58] S. Gupta and S. S. Sapatnekar, “Employing circadian rhythms to enhance power and reliability,” *ACM Transactions on Design Automation of Electronic Systems*, Vol. 18, No. 3, pp. 38:1–38:23, July 2013.
- [59] E. Cai, D. Stamoulis, and D. Marculescu, “Exploring aging deceleration in FinFET-based multi-core systems,” *Proceedings of the International Conference on Computer-Aided Design*, pp. 111:1–111:8, November 2016.
- [60] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, “The case for lifetime reliability-aware microprocessors,” *ACM SIGARCH Computer Architecture News*, Vol. 32, No. 2, pp. 276–288, June 2004.
- [61] A. S. Oates, “Reliability and technology scaling beyond the 10 nm node,” *Proceedings of the IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits*, pp. 1–3, June 2015.
- [62] G. Hubert, L. Artola, and D. Regis, “Impact of scaling on the soft error sensitivity of bulk, FDSOI and FinFET technologies due to atmospheric radiation,” *Integration, the VLSI journal*, Vol. 50, No. 1, pp. 39–47, June 2015.
- [63] K. V. Aadithya, A. Demir, S. Venugopalan, and J. Roychowdhury, “Accurate prediction of random telegraph noise effects in SRAMs and DRAMs,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 32, No. 1, pp. 73–86, January 2013.
- [64] Z. Li, M. Sotto, F. Liu, M. K. Husain, H. Yoshimoto, Y. Sasago, D. Hisamoto, I. Tomita, Y. Tsuchiya, and S. Saito, “Random telegraph noise from resonant tunnelling at low temperatures,” *Scientific reports*, Vol. 8, No. 1, pp. 250–259, January 2018.
- [65] X. Huang, W. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y. Choi, K. Asano, *et al.*, “Sub-50 nm P-channel FinFET,” *IEEE Transactions on Electron Devices*, Vol. 48, No. 5, pp. 880–886, May 2001.
- [66] Nanoscale Integration & Modeling Group Arizona State University, “Predictive Technology Model (PTM),” <http://www.eas.asu.edu/~ptm>.
- [67] W. Lee, Y. Wang, T. Cui, S. Nazarian, and M. Pedram, “Dynamic thermal management for FinFET-based circuits exploiting the temperature effect inversion phenomenon,” *Proceedings of the International Symposium on Low Power Electronics and Design*, pp. 105–110, August 2014.
- [68] S. E. Liu, J. S. Wang, Y. R. Lu, D. S. Huang, C. F. Huang, W. H. Hsieh, J. H. Lee, Y. S. Tsai, J. R. Shih, Y. H. Lee, and K. Wu, “Self-heating effect in

- FinFETs and its impact on devices reliability characterization,” *Proceedings of the Reliability Physics Symposium*, pp. 4A.4.1–4A.4.4, June 2014.
- [69] M. Shrivastava, M. Agrawal, J. Aghassi, H. Gossner, W. Molzer, T. Schulz, and V. R. Rao, “On the thermal failure in nanoscale devices: Insight towards heat transport including critical BEOL and design guidelines for robust thermal management and EOS/ESD reliability,” *Proceedings of the International Reliability Physics Symposium*, pp. 3F.3.1–3F.3.5, April 2011.
 - [70] H. H. Radamson, Y. Zhang, X. He, H. Cui, J. Li, J. Xiang, J. Liu, S. Gu, and G. Wang, “The challenges of advanced CMOS process from 2D to 3D,” *Applied Sciences*, Vol. 7, No. 10, pp. 1047–1079, October 2017.
 - [71] L. Wang, S. K. Khatamifard, U. R. Karpuzcu, and S. Köse, “Mitigation of NBTI induced performance degradation in on-chip digital LDOs,” *Proceedings of the Conference on Design, Automation and Test in Europe*, pp. 803–808, March 2018.
 - [72] D. Gnad, M. Shafique, F. Kriebel, S. Rehman, D. Sun, and J. Henkel, “Hayat: Harnessing dark silicon and variability for aging deceleration and balancing,” *Proceedings of the Design Automation Conference*, pp. 1–6, June 2015.
 - [73] A. F. Gomez and V. Champac, “Selection of critical paths for reliable frequency scaling under BTI-aging considering workload uncertainty and process variation effects,” *ACM Transactions on Design Automation of Electronic Systems*, Vol. 23, No. 3, pp. 27–48, February 2018.
 - [74] L. Benini, A. Bogliolo, and G.D. Micheli, “A survey of design techniques for system-level dynamic power management,” *Proceedings of the IEEE Transactions on Very Large Scale Integration Systems*, Vol. 8, No. 3, pp. 299–316, June 2000.
 - [75] S. S. Jha, W. Heirman, A. Falcón, T. E. Carlson, K. Van Craeynest, J. Tubella, A. González, and L. Eeckhout, “Chryso: An integrated power manager for constrained many-core processors,” pp. 19–26, May 2015.
 - [76] C. Isci, A. Buyuktosunoglu, C. Y. Cher, P. Bose, and M. Martonosi, “An analysis of efficient multi-core global power management policies: Maximizing performance for a given power budget,” pp. 347–358, December 2006.
 - [77] P. Hammarlund, A. J. Martinez, A. A. Bajwa, D. L. Hill, E. Hallnor, H. Jiang, M. Dixon, M. Derr, M. Hunsaker, R. Kumar, *et al.*, “Haswell: The fourth-generation intel core processor,” *Proceedings of the IEEE Micro*, Vol. 34, No. 2, pp. 6–20, March 2014.
 - [78] V. Subramaniana, B. Parvais, J. Borremans, A. Mercha, D. Linten, P. Wambacq, J. Loo, M. Dehan, N. Collaert, S. Kubicek, *et al.*, “Device and circuit-level analog performance trade-offs: a comparative study of planar bulk FETs versus FinFETs,” pp. 898–901, December 2005.

- [79] M. Borah, R. M. Owens, and M. J. Irwin, "Transistor sizing for low power CMOS circuits," *Proceedings of the IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 15, No. 6, pp. 665–671, June 1996.
- [80] V. Kursun and E. G. Friedman, "Low swing dual threshold voltage domino logic," *Proceedings of the ACM Great Lakes symposium on VLSI*, pp. 47–52, April 2002.
- [81] V. G. Oklobdzija, V. M. Stojanovic, D. M. Markovic, and N. M. Nedovic, *Digital system clocking: high-performance and low-power aspects*, John Wiley & Sons, 2005.
- [82] L. Wei, Z. Chen, K. Roy, M. C. Johnson, Y. Ye, and V. K. De, "Design and optimization of dual-threshold circuits for low-voltage low-power applications," *Proceedings of the IEEE Transactions on Very Large Scale Integration Systems*, Vol. 7, No. 1, pp. 16–24, March 1999.
- [83] A. Efthymiou and J. D. Garside, "Adaptive pipeline depth control for processor power-management," *Proceedings of the IEEE International Conference on Computer Design: VLSI in Computers and Processors*, pp. 454–457, September 2002.
- [84] D. Flynn, R. Aitken, A. Gibbons, and K. Shi, *Low power methodology manual: for system-on-chip design*, Springer Science & Business Media, 2007.
- [85] H. Soeleman, K. Roy, and B. Paul, "Sub-domino logic: ultra-low power dynamic sub-threshold digital logic," pp. 211–214, January 2001.
- [86] R. Zimmermann and W. Fichtner, "Low-power logic styles: CMOS versus pass-transistor logic," *IEEE journal of solid-state circuits*, Vol. 32, No. 7, pp. 1079–1090, 1997.
- [87] A. Buyuktosunoglu, D. Albonesi, S. Schuster, D. Brooks, P. Bose, and P. Cook, "A Circuit Level Implementation of an Adaptive Issue Queue for Power-aware Microprocessors," pp. 73–78, May 2001.
- [88] A. R. Alameldeen and D. A. Wood, "Adaptive cache compression for high-performance processors," *ACM SIGARCH Computer Architecture News*, Vol. 32, No. 2, pp. 212, 2004.
- [89] T. Li, D. Baumberger, D. A. Koufaty, and S. Hahn, "Efficient Operating System Scheduling for Performance-Asymmetric Multi-Core Architectures," *Proceedings of the ACM/IEEE conference on Supercomputing*, , No. 53, November 2007.
- [90] M. Pedram, D. Brooks, and T. Pinkston, "Report for the NSF workshop on cross-layer power optimization and management," July 2012.

- [91] I. Miro-Panades, E. Beigné, Y. Thonnart, L. Alacoque, P. Vivet, S. Lesecq, D. Puschini, A. Molnos, F. Thabet, B. Tain, *et al.*, “A fine-grain variation-aware dynamic Vdd-Hopping AVFS Architecture on a 32 nm GALS MPSoC,” *IEEE Journal of Solid-State Circuits*, Vol. 49, No. 7, pp. 1475–1486, July 2014.
- [92] T. Horvath, T. Abdelzaher, K. Skadron, and X. Liu, “Dynamic voltage scaling in multitier web servers with end-to-end delay control,” *IEEE Transactions on Computers*, Vol. 56, No. 4, pp. 444–458, April 2007.
- [93] M. Fojtik, D. Fick, Y. Kim, N. Pinckney, D. M. Harris, D. Blaauw, and D. Sylvester, “Bubble Razor: eliminating timing margins in an ARM cortex-M3 processor in 45 nm CMOS using architecturally independent error detection and correction,” *IEEE Journal of Solid-State Circuits*, Vol. 48, No. 1, pp. 66–81, January 2013.
- [94] D. Ernst, N. S. Kim, S. Das, S. Pant, R. Rao, T. Pham, C. Ziesler, D. Blaauw, T. Austin, K. Flautner, *et al.*, “Razor: A low-power pipeline based on circuit-level timing speculation,” pp. 7–18, December 2003.
- [95] M. Tavana, D. Pathak, M. Hajkazemi, I. Savidis, and H. Homayoun, “Realizing complexity-effective on-chip power delivery for many-core platforms by exploiting optimized mapping,” *Proceedings of the IEEE International Conference on Computer Design*, pp. 581–588, October 2015.
- [96] M. K. Tavana, M. H. Hajkazemi, D. Pathak, I. Savidis, and H. Homayoun, “ElasticCore: A dynamic heterogeneous platform with joint core and voltage/frequency scaling,” *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 26, No. 2, pp. 249–261, Feb 2018.
- [97] mersenne.org, “Great Internet Mersenne Prime Search,” <http://www.mersenne.org/download/#stresstest>.
- [98] S. Li, J.H. Ahn, R.D. Strong, J.B. Brockman, D.M. Tullsen, and N.P. Jouppi, “McPAT: an integrated power, area, and timing modeling framework for multi-core and manycore architectures,” *Proceedings of the IEEE/ACM International Symposium on Microarchitecture*, pp. 469–480, October 2009.
- [99] W. Godycki, C. Torng, I. Bukreyev, A. Apsel, and C. Batten, “Enabling realistic fine-grain voltage scaling with reconfigurable power distribution networks,” *Proceedings of the IEEE/ACM International Symposium on Microarchitecture*, pp. 381–393, December 2014.
- [100] W. Lee, Y. Wang, and M. Pedram, “Optimizing a reconfigurable power distribution network in a multicore platform,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, January 2015.
- [101] M. Sai and H. Yu, “3D Many-core Microprocessor Power Management by Space-Time Multiplexing based Demand-supply Matching,” *IEEE Transaction on Computers*, Vol. 64, No. 11, pp. 3022–3036, January 2015.

- [102] I. Vaisband and E. G. Friedman, “Energy efficient adaptive clustering of on-chip power delivery systems,” *INTEGRATION, the VLSI Journal*, Vol. 48, pp. 1–9, January 2015.
- [103] T. N. Miller, X. Pan, R. Thomas, N. Sedaghati, and R. Teodorescu, “Booster: Reactive Core Acceleration for Mitigating the Effects of Process Variation and Application Imbalance in Low-Voltage Chips,” *Proceedings of the International Symposium on High Performance Computer Architecture*, pp. 1–12, February 2012.
- [104] D. M. Tullsen, “Simulation and modeling of a simultaneous multithreading processor,” *Proceedings of the International Conference for the Resource Management & Performance Evaluation of Enterprise Computing Systems, CMG. Part 2(of 2)*, pp. 819–828, June 1996.
- [105] K. Pearson, “Contributions to the Mathematical Theory of Evolution,” *Philosophical Transactions of the Royal Society of London*, Vol. 185, pp. 71–110, January 1894.
- [106] Texas Instruments, “WEBENCH Design Center,” <http://webench.ti.com>.
- [107] Y. Choi, N. Chang, and T. Kim, “DC–DC Converter-Aware Power Management for Low-Power Embedded Systems,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 26, No. 8, pp. 1367–1381, August 2007.
- [108] H. C. Cranford Jr, L. L. C. Hsu, and J. S. Mason, “Power network reconfiguration using MEM switches,” December 2007, US Patent 7,305,571.
- [109] O. Y. Loh and H. D. Espinosa, “Nanoelectromechanical Contact Switches,” *Nature Nanotechnology*, Vol. 7, No. 5, pp. 283–295, April 2012.
- [110] Z. Li, R. Balasubramanian, F. Liu, and S. Nassif, “2012 TAU Power Grid Simulation Contest: Benchmark Suite and Results,” *Proceedings of the ACM International Conference on Computer-Aided Design*, pp. 643–646, November 2012.
- [111] IBM, “IBM Power Grid Benchmarks,” <http://dropzone.tamu.edu/~pli/PGBench/>.
- [112] Texas Instruments, “DC to DC Converter Parallel Operation Using Droop Compensation of TPS5210,” <http://www.ti.com/lit/an/slva060/slva060.pdf>.
- [113] D. Pathak, H. Homayoun, and I. Savidis, “Work load scheduling for multi core systems with under-provisioned power delivery,” *Proceedings of the on Great Lakes Symposium on VLSI 2017*, pp. 387–392, May 2017.

- [114] S. Ghiasi, T. Keller, and F. Rawson, "Scheduling for Heterogeneous Processors in Server Systems," *Proceedings of the Conference on Computing Frontiers*, pp. 199–210, May 2005.
- [115] G. Buttazzo, *Hard real-time computing systems: predictable scheduling algorithms and applications*, Vol. 24, Springer Science & Business Media, 2011.
- [116] H. Chung, M. Kang, and H. Cho, "Heterogeneous Multi-Processing Solution of Exynos 5 Octa with ARM® big.LITTLE Technology," *Samsung White Paper*, 2012.
- [117] Y. Shin, H.J. Lee, K. Shin, P. Kenkae, R. Kashyap, D. Seo, B. Millar, Y. Kwon, R. Iyengar, M. Kim, *et al.*, "28nm High-K Metal Gate Heterogeneous Quad-Core CPUs for High-Performance and Energy-Efficient Mobile Application Processor," *Proceedings of the IEEE International SoC Design Conference*, pp. 198–201, November 2013.
- [118] A. Colin, A. Kandhalu, and R. Rajkumar, "Energy-efficient allocation of real-time applications onto heterogeneous processors," *Proceedings of the IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, pp. 1–10, August 2014.
- [119] V. J. Reddi, S. Kanev, Kim. W., S. Campanoni, M. D. Smith, G. Y. Wei, and D. Brooks, "Voltage Smoothing: Characterizing and mitigating voltage noise in production processors via software-guided thread scheduling," *IEEE/ACM International Symposium on Microarchitecture*, pp. 77–88, December 2010.
- [120] ITRS, "International Technology Roadmap for Semiconductors," <http://www.itrs2.net/>.
- [121] M. Popovich, A. Mezhiba, and E. G. Friedman, *Power distribution networks with on-chip decoupling capacitors*, Springer Science & Business Media, 2007.
- [122] S. Kose and E.G. Friedman, "On-chip point-of-load voltage regulator for distributed power supplies," *Proceedings of the IEEE Great lakes symposium on VLSI*, pp. 377–380, 2010.
- [123] T. Charania, A. Opal, and M. Sachdev, "Analysis and design of on-chip decoupling capacitors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 21, No. 4, pp. 648–658, July 2013.
- [124] A. Drake, R. Senger, H. Deogun, G. Carpenter, S. Ghiasi, T. Nguyen, N. James, M. Floyd, and V. Pokala, "A distributed critical-path timing monitor for a 65nm high-performance microprocessor," pp. 398–399, February 2007.
- [125] P. N. Whatmough, S. Das, and D. M. Bull, "Analysis of adaptive clocking technique for resonant supply voltage noise mitigation," *International Symposium on Low Power Electronics and Design*, pp. 128–133, July 2015.

- [126] D. Pathak and I. Savidis, “On-Chip power supply noise suppression through hyperabrupt junction varactors,” *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 26, No. 11, pp. 2230–2240, November 2018.
- [127] M. D Pant, P. Pant, and D. S. Wills, “On-chip decoupling capacitor optimization using architectural level prediction,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 10, No. 3, pp. 319–326, June 2002.
- [128] D. Stringfellow and J. Pedicone, “Decoupling capacitance estimation, implementation, and verification: A practical approach for deep submicron SoCs,” *Synopsys Users Group*, 2007.
- [129] L. D. Smith and E. Bogatin, *Principles of power integrity for PDN design: Robust and cost effective design for high speed digital products*, Prentice Hall, 2017.
- [130] P. Penfield and R. P. Rafuse, *Varactor applications*, MIT Press, 1962.
- [131] J. Heer, “Surface barrier diode having a hypersensitive n region forming a hypersensitive voltage variable capacitor,” December 1972, US Patent 3,706,128.
- [132] D. Pavlidis, Y. Archambault, and L. Karapuperis, “Variable capacity diode with hyperabrupt profile and plane structure and the method of forming same,” May 1989, US Patent 4,827,319.
- [133] C. A. Goodwin and Y. Ota, “Method of making a hyperabrupt varactor diode utilizing molecular beam epitaxy,” October 1980, US Patent 4,226,648.
- [134] M. N. Yoder, “Method of forming a hyperabrupt interface in a GaAs substrate,” July 1983, US Patent 4,391,651.
- [135] J. B. Johnson, A. J. Joseph, R. M. Rassel, and Y. Shi, “Device structures with a hyper-abrupt PN junction, methods of forming a hyper-abrupt PN junction, and design structures for an integrated circuit,” September 2010, US Patent 7,804,119.
- [136] Skyworks Solutions Inc, “Varactor diodes application note,” <http://skyworksinc.com/uploads/documents/200824A.pdf>.
- [137] R. Cory, “The Nuts and Bolts of Tuning Varactors,” *High Frequency Electronics*, pp. 42–51, February 2009.
- [138] Skyworks Solutions Inc., “SMVA1211-001LF: Hyperabrupt Junction Tuning Varactor,” <http://skyworksinc.com/Product/1799/SMVA1211-001LF>.
- [139] M. A. Ang and A. D. Taylor, “Voltage regulator circuit for attenuating inductance-induced on-chip supply variations,” February 2000, US Patent 6,028,417.

- [140] A. D. Taylor and M. A. Ang, “Voltage regulation method for attenuating inductance-induced on-chip supply variations,” May 2000, US Patent 6,069,521.
- [141] D. T. Blaauw, R. V. Panda, R. Chaudhry, V. P. Zolotov, and R. Ramaraju, “Method and apparatus for controlling current demand in an integrated circuit,” November 2004, US Patent 6,819,538.
- [142] Q. Li, “Power distribution network (pdn) conditioner,” February 2014, US Patent 14/175,922.
- [143] W. Abadeer, R. Rassel, and J. B. Johnson, “A capacitance reliability degradation mechanism in Hyper-abrupt junction varactors,” *IEEE International Reliability Physics Symposium*, pp. 310–314, April 2008.
- [144] R. W. Berry, M. S. Floyd, J. Pena, R. J. Pennington, and C. Sherry, “Processor noise mitigation using differential critical path monitoring,” October 2015, US Patent 9,164,563.
- [145] M. C. Pedrone, K. D. Peterson, E. S. I. I. John, and A. A. Turner, “Dynamic noise mitigation in integrated circuit devices using local clock buffers,” July 2017, US Patent 9,712,112.
- [146] M. S. Floyd, P. J. Restle, M. A. Sperling, P. Owczarczyk, E. J. Fluhr, J. Friedrich, P. Muench, T. Diemoz, P. Chuang, and C. Vezirtzis, “Adaptive clocking in the POWER9 processor for voltage droop protection,” *IEEE International Solid-State Circuits Conference*, pp. 444–445, February 2017.
- [147] D. A. Kamakshi, M. Fojtik, B. Khailany, S. Kudva, Y. Zhou, and B. H. Calhoun, “Modeling and Analysis of Power Supply Noise Tolerance with Fine-Grained GALS Adaptive Clocks,” *IEEE International Symposium on Asynchronous Circuits and Systems*, pp. 75–82, May 2016.
- [148] Y. Koda, H. Sugita, T. Suwa, R. Kuroda, T. Goto, A. Teramoto, and S. Sugawa, “Low leakage current Al₂O₃ metal-insulator-metal capacitors formed by atomic layer deposition at optimized process temperature and O₂ post deposition annealing,” *ECS Transactions*, Vol. 72, No. 4, pp. 91–100, May 2016.
- [149] R. Aparicio and A. Hajimiri, “Capacity limits and matching properties of integrated capacitors,” *IEEE Journal of Solid-State Circuits*, Vol. 37, No. 3, pp. 384–393, August 2002.
- [150] Q. S. Lim, A. V. Kordesch, and R. A. Keating, “Performance comparison of MIM capacitors and metal finger capacitors for analog and RF applications,” *Proceedings of the RF and Microwave Conference*, pp. 85–89, April 2004.
- [151] J. H. Klootwijk, K. B. Jinesh, W. Dekkers, J. F. Verhoeven, F. C. Heuvel, H. D. Kim, D. Blin, M. A. Verheijen, R. G. R. Weemaes, M. Kaiser, *et al.*, “Ultrahigh capacitance density for multiple ALD-grown MIM capacitor stacks

- in 3-D silicon,” *IEEE Electron Device Letters*, Vol. 29, No. 7, pp. 740–742, July 2008.
- [152] G. Wang, D. Anand, N. Butt, A. Cestero, M. Chudzik, J. Ervin, S. Fang, G. Freeman, H. Ho, B. Khan, *et al.*, “Scaling deep trench based eDRAM on SOI to 32nm and Beyond,” *IEEE International Electron Devices Meeting*, pp. 1–4, December 2009.
 - [153] I. Savidis, *Characterization and modeling of TSV based 3-D integrated circuits*, Ph.D. Thesis, University of Rochester, 2013.
 - [154] D. Pathak and I. Savidis, “Run-time voltage detection circuit for 3-D IC power delivery,” *Proceedings of the IEEE International SoC Conference*, pp. 229–233, September 2014.
 - [155] D. Pathak and I. Savidis, “Power supply voltage detection and clamping circuit for 3-D integrated circuits,” *Proceedings of the IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference*, pp. 1–3, October 2014.
 - [156] I. Savidis and D. Pathak, “Power supply voltage detection and power delivery circuit,” October 2016, US Patent 9912325.
 - [157] F. Kashfi and J. Draper, “Thermal sensor design for 3D ICs,” *Proceedings of the IEEE International Midwest Symposium on Circuits and Systems*, pp. 482–485, August 2012.
 - [158] S. Lopez-Buedo, J. Garrido, and E. Boemo, “Thermal testing on reconfigurable computers,” *IEEE Design & Test of Computers*, Vol. 17, No. 1, pp. 84–91, March 2000.
 - [159] J. S. Panganiban, *A ring oscillator based variation test chip*, Master’s thesis, Massachusetts Institute of Technology, 2002.
 - [160] G. Jovanovic, M. Stojcev, and Z. Stamenkovic, “A CMOS voltage controlled ring oscillator with improved frequency stability,” *Scientific Publications of the State University of Novi Pazar, Series A: Applied Mathematics, Informatics and Mechanics*, Vol. 2, No. 1, pp. 1–9, December 2010.
 - [161] A. Ramazani, S. Biabani, and G. Hadidi, “CMOS ring oscillator with combined delay stages,” *AEU-International Journal of Electronics and Communications*, Vol. 68, pp. 515–519, May 2014.
 - [162] X. Zhang, A. M. Pappu, and A.B. Apsel, “Low variation current source for 90nm CMOS,” *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 388–391, May 2008.
 - [163] X. Zhang and A.B. Apsel, “A low-power, process-and-temperature-compensated ring oscillator with addition-based current source,” *IEEE Transactions on Circuits and Systems*, Vol. 58, No. 5, pp. 868–878, May 2011.

- [164] Y. Cao, *Predictive Technology Model for Robust Nanoelectronic Design*, Springer, 2011.
- [165] S. Kose and E.G. Friedman, “Distributed on-chip power delivery,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, Vol. 2, No. 4, pp. 704–713, December 2012.
- [166] A. Djemouai, M. Sawan, and M. Slamani, “High performance integrated CMOS frequency-to-voltage converter,” *Proceedings of the Tenth International Conference on Microelectronics*, pp. 63–66, December 1998.
- [167] North Carolina State University, “FreePDK45,” <http://www.eda.ncsu.edu/wiki/FreePDK>.
- [168] I. Savidis and E. G. Friedman, “Electrical modeling and characterization of 3-D vias,” *IEEE International Symposium on Circuits and Systems*, pp. 784–787, May 2008.
- [169] I. Savidis, S. M. Alam, A. Jain, S. Pozder, R. E. Jones, and R. Chatterjee, “Electrical modeling and characterization of through-silicon vias (TSVs) for 3-D integrated circuits,” *Microelectronics Journal*, Vol. 41, No. 1, pp. 9–16, January 2010.
- [170] I. Savidis and E. G. Friedman, “Closed-form expressions of 3-D via resistance, inductance, and capacitance,” *IEEE Transactions on Electron Devices*, Vol. 56, No. 9, pp. 1873–1881, September 2009.
- [171] J.R. Yuan and C. Svensson, “Fast CMOS nonbinary divider and counter,” *Electronics Letters*, Vol. 29, No. 13, pp. 1222–1223, June 1993.
- [172] R.J. Baker, *CMOS: circuit design, layout, and simulation*, Vol. 18, John Wiley & Sons, 2011.
- [173] Analog Devices, “ADP1821 Step-Down DC-to-DC Controller Datasheet,” http://www.analog.com/static/imported-files/data_sheets/ADP1821.pdf.
- [174] X. Jiao, M. Luo, J. Lin, and R. K. Gupta, “An assessment of vulnerability of hardware neural networks to dynamic voltage and temperature variations,” *Proceedings of the International Conference on Computer-Aided Design*, pp. 945–950, November 2017.
- [175] J. Zhang, K. Rangineni, Z. Ghodsi, and S. Garg, “ThUnderVolt: Enabling aggressive voltage undervolting and timing error resilience for energy efficient deep neural network accelerators,” *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 1–6, June 2018.
- [176] Y. Zu, C. R. Lefurgy, J. Leng, M. Halpern, M. S. Floyd, and V. J. Reddi, “Adaptive guardband scheduling to improve system-level efficiency of the POWER7+,” *Proceedings of the International Symposium on Microarchitecture*, pp. 308–321, December 2015.

- [177] Sanjay Pant, *Design and analysis of power distribution networks in VLSI circuits.*, Ph.D. Thesis, University of Michigan, 2008.
- [178] W. H. Chang, C. H. Lin, S. P. Mu, L. D. Chen, C. H. Tsai, Y. C. Chiu, and M. Chao, “Generating routing-driven power distribution networks with machine-learning technique,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 36, No. 8, pp. 1237–1250, January 2017.
- [179] F. Ye, F. Firouzi, Y. Yang, K Chakrabarty, and M. B. Tahoori, “On-chip droop-induced circuit delay prediction based on support-vector machines,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 35, No. 4, pp. 665–678, April 2016.
- [180] A. P. Da Silva and P. Abrão, “Applications of evolutionary computation in electric power systems,” *Proceedings of the Congress on Evolutionary Computation*, Vol. 2, pp. 1057–1062, May 2002.
- [181] N. Guofang, L. Minqiang, and K. Jisong, “Multi-level genetic algorithm (MLGA) for the construction of clock binary tree,” *Proceedings of the Conference on Genetic and Evolutionary Computation*, pp. 1441–1445, June 2005.
- [182] T. Susa, M. Murakawa, E. Takahashi, T. Furuya, and T. Higuchi, “Post-fabrication clock-timing adjustment for digital LSIs ensuring operational timing margins,” *Proceedings of the Hybrid Intelligent Systems conference*, pp. 907–910, March 2008.
- [183] P. N. Whatmough, S. Das, Z. Hadjilambrou, and D. M. Bull, “Power integrity analysis of a 28 nm dual-core ARM Cortex-A57 cluster using an all-digital power delivery monitor,” *IEEE Journal of Solid-State Circuits*, Vol. 52, No. 6, pp. 1643–1654, June 2017.
- [184] D. W. Boeringer and D. H. Werner, “Particle swarm optimization versus genetic algorithms for phased array synthesis,” *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 3, pp. 771–779, April 2004.
- [185] L. Zhang, Y. Chen, R. Sun, S. Jing, and B. Yang, “A task scheduling algorithm based on PSO for grid computing,” *International Journal of Computational Intelligence Research*, Vol. 4, No. 1, pp. 37–43, June 2008.
- [186] Y. Shi and R. C. Eberhart, “Parameter selection in particle swarm optimization,” *Proceedings of the International Conference on Evolutionary Programming*, pp. 591–600, March 1998.
- [187] R. Franch, P. Restle, N. James, W. Huott, J. Friedrich, R. Dixon, S. Weitzel, K. Van Goor, and G. Salem, “On-chip timing uncertainty measurements on IBM microprocessors,” *Proceedings of the IEEE International Test Conference*, pp. 1–7, October 2007.

- [188] M. S. Floyd, A. Drake, N. S. Schwartz, R. W. Berry, C. R. Lefurgy, M. Ware, K. Rajamani, V. Zyuban, R. Willaman, and R. M. Zgabay, “Runtime power reduction capability of the IBM POWER7+ chip,” *IBM Journal of Research and Development*, Vol. 57, No. 6, pp. 1–17, November 2013.
- [189] J. Park and J. A. Abraham, “A fast, accurate and simple critical path monitor for improving energy-delay product in DVS systems,” *Proceedings of the International Symposium on Low-power Electronics and Design*, pp. 391–396, August 2011.
- [190] M. Agarwal, B. C. Paul, M. Zhang, and S. Mitra, “Circuit failure prediction and its application to transistor aging,” *Proceedings of the IEEE VLSI Test Symposium*, pp. 277–286, May 2007.
- [191] R. Zhang, K. Mazumdar, B. H. Meyer, K. Wang, K. Skadron, and M. R. Stan, “Voltspot 2.0,” <http://lava.cs.virginia.edu/VoltSpot/>.
- [192] R. Zhang, K. Skadron, and M. R. Stan, “Hotspot 6.0,” <http://lava.cs.virginia.edu/HotSpot/>.
- [193] J. Leng, Y. Zu, M. Rhu, M. Gupta, and V. J. Reddi, “GPUVolt: Modeling and characterizing voltage noise in GPU architectures,” *Proceedings of the International Symposium on Low Power Electronics and Design*, pp. 141–146, August 2014.
- [194] P. I. Chuang, C. Vezirtzis, D. Pathak, R. Rizzolo, T. Webel, T. Strach, O. Torreiter, P. Lobo, A. Buyuktosunoglu, R. Bertran, M. Floyd, M. Ware, G. Salem, S. Carey, and P. Restle, “Power supply noise in a 22nm z13™ microprocessor,” *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 438–439, February 2017.
- [195] J. L. Hennessy and D. A. Patterson, *Computer architecture: a quantitative approach*, Elsevier, 2018.
- [196] K. M. Zick, M. Srivastav, W. Zhang, and M. French, “Sensing nanosecond-scale voltage attacks and natural transients in FPGAs,” *Proceedings of the ACM/SIGDA International Symposium on Field Programmable Gate Arrays*, pp. 101–104, 2013.

Vita

Divya Pathak received the B.Tech degree in electronics and communication engineering from IGIT, Delhi, India in 2002 and M.S. degree in computer engineering from Drexel University, Philadelphia, USA, in 2015 where she is expected to complete the Ph.D. degree in computer engineering in 2018.



She was a Research Scientist with Space Applications Center, Indian Space Research Organization, from 2003 to 2006, where she designed and developed surface acoustic wave devices for analog signal processing in communication and remote sensing satellites. She was a Staff Engineer with STMicroelectronics from 2006 to 2013, where she was the lead in functional and electrical post-silicon validation of high speed SERDES links in MPSoCs. She was a Research Intern with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA, in 2015 and 2016, where she was involved in power supply noise mitigation techniques for IBM POWER and Z Systems.

Academic Honors

1. Awarded the 2017 IEEE Circuits and Systems Society Pre-Doctoral Scholarship.
2. Awarded the 2017 National Science Foundation (NSF) Professional Development Award as part of the iREDEFINE program (NSF Grant number 1663249).

3. Recipient of the Joseph and Shirley Carleone Endowed Fellowship, College of Engineering, Drexel University, 2017-2018.
4. Recipient of the Allen Rothwarf Scholarship, Department of Electrical and Computer Engineering, Drexel University, 2016-2017.
5. Awarded the 2016 IEEE Circuits and Systems Society Student Travel Award and the International Travel Award from Drexel University to present paper at ISCAS 2016 in Montreal, Canada.
6. Recipient of the Seaman Fellowship, Department of Electrical and Computer Engineering, Drexel University, 2014-2015.

Journal Publications

5. **D. Pathak** and I. Savidis, "Evolving On-Chip Power Delivery through Particle Swarm Optimization," *IEEE Transactions on Circuits and Systems I (TCAS-I)*, [submitted].
4. **D. Pathak** and I. Savidis, "On-Chip Power Supply Noise Suppression Through Hyperabrupt Junction Varactors," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems*, Vol. 26, no. 11, pp. 2230-2240, November 2018.
3. M. Tavana, M. Hajkazemi, **D. Pathak**, I. Savidis, and H. Homayoun, "ElasticCore: A Dynamic Heterogeneous Platform with Joint Core and Voltage/Frequency Scaling," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems*, Vol. 26, no. 9, pp. 249-261, January 2018.

2. **D. Pathak**, H. Homayoun, and I. Savidis, "SMART GRID ON CHIP: Load Balanced On-Chip Power Delivery," *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems*, Vol. 25, no. 9, pp. 2538-2551, September 2017.
1. V. Sinha, **D. Pathak**, A. Kumar, and O.P. Kaushik, "Survey of SAW Devices in Space," *International Journal of Space Craft technology*, Vol. 18, no. 2, pp. 61-75, July 2008.

Conference Publications

13. **D. Pathak** and I. Savidis. "Applying swarm intelligence to distributed on-chip power management," *Submitted to ACM Design Automation Conference 2019* (DAC).
12. H. Sayadi, **D. Pathak**, I. Savidis, and H. Homayoun "Power Conversion Efficiency-Aware Mapping of Multithreaded Applications on Heterogeneous Architectures: A Comprehensive Parameter Tuning", *Proceedings of the Asia and South Pacific Design Automation Conference (ASPDAC)*, January 2018.
11. **D. Pathak**, H. Homayoun, and I. Savidis, "Work Load Scheduling For Multi Core Systems With Under-Provisioned Power Delivery", *Proceedings of the IEEE/ACM Great Lakes Symposium on VLSI (GLSVLSI)*, pp. 387–392, May 2017.
10. P. Chuang, C. Vezyrtzis, **D. Pathak**, R. Rizzolo, T. Webel, T. Strach, O.Torreiter, P. Lobo, A. Buyuktosunoglu, R. Bertran, M. Floyd, M. Ware, G. Salem, S.

- Carey, P. Restle, "Power Supply Noise in a 22nm $z13^{TM}$ Microprocessor," *Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 438–439, February 2017.
9. **D. Pathak**, M. Hajkazemi, M. Tavana, H. Homayoun, and I. Savidis, "Load Balanced On-Chip Power Delivery for Average Current Demand", *Proceedings of the IEEE/ACM Great Lakes Symposium on VLSI (GLSVLSI)*, pp. 439–444, May 2016.
 8. **D. Pathak**, M. Hajkazemi, M. Tavana, H. Homayoun, and I. Savidis, "Energy Efficient On-Chip Power Delivery with Run-Time Voltage Regulator Clustering", *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1210–1213, May 2016.
 7. M. Tavana, **D. Pathak**, M. Hajkazemi, M. Malik, I. Savidis, and H. Homayoun, "Realizing complexity-effective on-chip power delivery for many-core platforms by exploiting optimized mapping," *Proceedings of the IEEE International Conference on Computer Design (ICCD)*, pp. 581–588, October 2015.
 6. M. Tavana, M. Hajkazemi, **D. Pathak**, I. Savidis, and H. Homayoun, "ElasticCore: enabling dynamic heterogeneity with joint core and voltage/frequency scaling," *Proceedings of the 52nd ACM Design Automation Conference (DAC)*, pp. 151–157, June 2015.
 5. **D. Pathak** and I. Savidis. "Run-Time Voltage Detection Circuit for 3-D IC Power Delivery," *Proceedings of the IEEE 27th International SoC Conference*

(SoCC), pp. 183–187, September 2014.

4. **D. Pathak** and I. Savidis, "Power Supply Voltage Detection and Clamping Circuit for 3-D Integrated Circuits," *Proceedings of the IEEE SOI-3D- Subthreshold Microelectronics Technology Unified Conference (S3S)*, pp. 94–96, October 2014.
3. **D. Pathak** and A. Kumar, "Design and Development of Wide Band Linear Phase Slanted IDT SAW Filter," *presented at 3rd International Conference on Microwave, Antenna, Propagation and Remote sensing*, December 2006.
2. A. Kumar and **D. Pathak**, "SAW Devices for GAGAN Payload," presented at ISRO Intercentre technical seminar, November 2005, ISRO Satellite Centre, India.
1. J. P. Pabari, H. Dave, V. Sinha, D. Balasubramaniam, A. B. Shah, A. Dubey, N. M. Vadher, V. Shah, S. Thampi, R. P. Singh, V. Kumar, A. Kumar, and **D. Pathak**, "Chirp Transform Spectrometer Development for Submillimeter Wave Spectroscopy," *Proceedings of International Conference on Sub millimeter Science and Technology (ICSST)*, pp. 82–85, October 2004.

Patents

5. I. Savidis and **D. Pathak**, "Power Supply Voltage Detection and Power Delivery Circuit," *Patent grant, U.S. Patent No. 9,912,325*.
4. P. Chuang, **D. Pathak**, P. Restle, and, C. Vezyrtzis, "Mitigation of on-chip

supply voltage noise by monitoring slope of supply voltage based on time-based sensors," *Patent pending, U.S. Patent Application No. 2018/0067532*.

3. I. Savidis, **D. Pathak**, and, H. Homayoun, "Work load scheduling for multi core systems with underprovisioned power delivery," *Patent pending, US Patent Application No. 15/968,348*.
2. **D. Pathak** and I. Savidis, "On-Chip Power Supply Noise Suppression Through Hyperabrupt Junction Varactors," *Patent pending, U.S. Patent Application No. 62,688,022*.
1. **D. Pathak** and I. Savidis, "On-Chip Voltage Assignment through Particle Swarm Optimization," *Patent pending, U.S. Patent Application No. 62,726,345*.

