Circuit Design and Modeling Techniques of On-chip Power Delivery Modules

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

AYAN PAUL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

CHRIS H. KIM

January 2015

Acknowledgements

Firstly, I would like to thank Prof. Chris H. Kim, my PhD advisor. I am indebted to him for his guidance during my Ph.D. at the University of Minnesota.

Secondly, I would like to thank my Ph.D. committee: Prof. Sachin S. Sapatnekar, Prof. Steven J. Koester and Prof. Antonia Zhai. Their valuable comments and suggestions helped me improve my thesis.

Last but not least, I would like to thank all my colleagues in the VLSI Research Group in the University of Minnesota. They are: Dr. Tony Kim, Dr. John Keane, Dr. Kichul Chun, Dr. Wei Zhang, Dr. Pulkit Jain, Dr. Xiaofei Wang, Dr. Seung-hwan Song, Dr. Bongjin Kim, Jongyeon Kim, Won Ho Choi, Saroj Satapathy, Weichao Xu, Qianying Tang, Somnath Kundu, Chen Zhou, Saurabh Kumar, Dr. Hoonki Kim, and Ibrahim Ahmed.

Dedication

To my friends and family

Abstract

Designing an efficient power delivery network is one of the most important aspects of modern low-power microprocessor designs. Power delivery network consists of off-chip converters, wires/bumps to route power and ground signals from off-chip converters to inside the chip, on-chip converters, and on-chip power grid. Parasitic resistance and inductance of routing wires and package cause IR noise and resonant supply noise in the supply line, whereas parasitic capacitance of the power and ground lines results in a significant increase in the charging and discharging time of the supply lines during transients of dynamic voltage and frequency scaling (DVFS). DVFS, which scales supply voltage depending on the performance requirement of the processor, thereby saving dynamic power dissipation, has become an integral part of today's microprocessor. However, because of large IR noise to route off-chip power and ground signals and large off-chip component count, per-core DVFS in a many-core processor cannot be supported with off-chip converters.

We present design techniques to tackle non-idealities of the power delivery networks. At the same time, we propose on-chip power delivery solutions. We present a circuit technique based on staggering activation of cores in order to mitigate first-droop noise from the supply network. From our 65nm bulk CMOS test-chip, we measure a 12.7% improvement in resonant supply noise of a 2-core processor using our proposed noise reduction technique. In order to reduce IR noise of the power delivery network, we propose a switched capacitor step-up converter. The converter built in 32nm SOI CMOS process shows a 45% improvement in IR noise of the supply line. For faster charging of the supply line, we build a step-down converter in 32nm SOI CMOS process, and measured 5x reduction in the charging time of the supply network. Using a 32nm SOI CMOS process, we build an on-chip switched capacitor DC/DC converter, capable of supporting two outputs with the help of time division demultiplexing. Finally, we propose two power delivery system using two industry-standard state-of-the-art power delivery units (fully

integrated voltage regulators and low dropout regulators), and compare the power consumption of the entire system using random and minimum power scheduling techniques.

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

Increased transistor counts and improved frequency of operation with transistor scaling necessitate implementation of various low power techniques along with improvement in the design of the heat sink and the packaging technology. Many low power design techniques such as dynamic voltage and frequency scaling (DVFS), clock gating, power gating etc. have been proposed to combat the issue of high power density associated with MOSFET scaling. Clock gating technique reduces clock power by selectively powering down the clock signal of idle blocks inside the cores [1]. Power gating technique uses a sleep transistor/ power gate that connects local supply line of a processor block to the global supply [2]. By turning the sleep

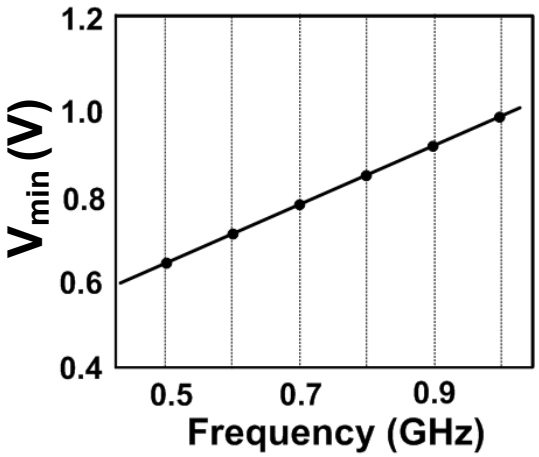


Figure 1.1: V_{min} scales linearly with frequency.

transistor off of an idle block, power gating technique allows to save leakage power dissipated by the cores.

The idea of DVFS was first presented in the 90's [3]. In order to run the cores at a particular frequency (f), the core supply voltage (V_{DD}) can be reduced to a certain minimum value, V_{min}, below which the cores will cease to maintain its operating frequency, f. Dynamic power

consumed by the cores is given by $P_{dyn} = \alpha \cdot C_{EFF} \cdot V_{DD}^2 . f$, where α is the activity factor, C_{EFF} is the effective capacitance of the core. The expression for P_{dyn} indicates that dynamic power can scale in a linear fashion with f, but in a quadratic fashion with V_{DD}. Since V_{DD} can be scaled down to V_{min} to sustain a frequency of f, it allows room for reduction of dynamic power by scaling V_{DD} down to V_{min}. This technique of scaling V_{DD} based on operating frequency is known as DVFS. In deep-submicron technologies, V_{DD} scales linearly with operating frequency, f. Figure 1.1 plots V_{DD} as a function of f. Because of this linear dependence between V_{DD} and f, it is possible to scale down dynamic power consumptions of the cores in a cubic fashion as f reduces linearly. This can be seen from Figure 1.2 , in which the blue line indicates that without DVFS, dynamic power reduces linearly as operating frequency scales down. However, with DVFS, dynamic power can be reduced more aggressively for linear reduction in operating frequency, as shown by the green line in Figure 1.2.

The effectiveness of these low power techniques is dependent on the sizes of the blocks that they

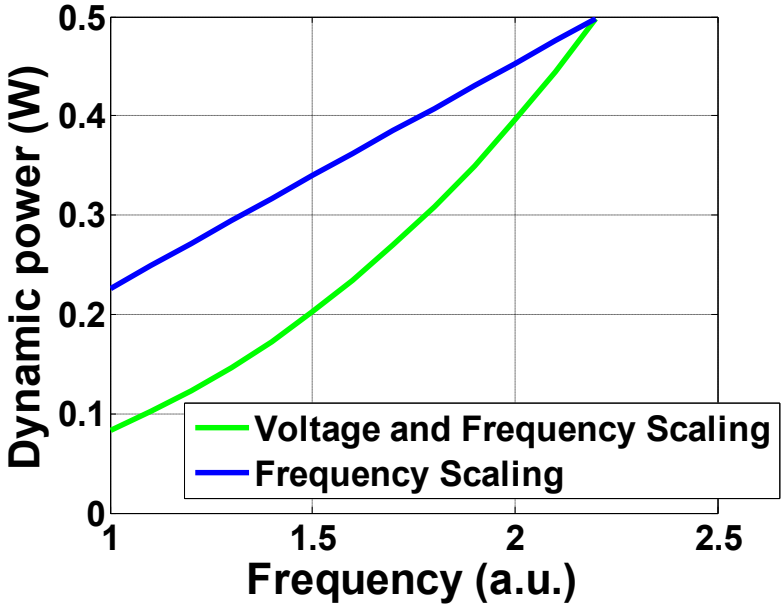


Figure 1.2: Dynamic power vs. frequency without and with DVFS.

are controlling. For example, DVFS, when applied on a per-core basis within a microprocessor, can result in maximum reduction in power dissipation, as opposed to when it is applied on a cluster of cores. This is because of the fact that in the latter case more than one cores share a single power domain. Hence, the supply voltage of that domain is determined by the supply voltage requirement of the core operating with highest frequency, thereby rendering this configuration sub-optimal for the cores running at lower frequencies.

Although per-core DVFS results in maximum reduction in power dissipated by the cores, its implementation in a many-core processor becomes unrealistic when the number of cores in the microprocessor increases because of the limitations associated with the design of power delivery networks. Thus, proper designs of power delivery networks and modules play a major role of low power circuit designs.

There are various non-idealities associated with the power delivery networks. Power delivery networks consist of

1. off-chip voltage regulators, which convert battery voltage or the voltage coming from the main supply to the voltage level that can be used by on-chip voltage regulators or microprocessor cores
2. on-chip voltage regulators, which take the output voltage of the off-chip regulator and generate voltage that will be used by microprocessor cores
3. wires/bumps to route the power signals from the off-chip converter to inside the chip
4. on-chip power grid for on-chip routing of power signals.

Because of the presence of larger parasitic resistance in the routing network, there is a significant IR drop in the supply voltage level. On the other hand, parasitic inductance associated with motherboard socket, package metal and long metal lines of the on-chip routing network results in large spikes in the supply voltage during core wake-up and core sleep transients. Large noise present in the supply line limits the performance of the cores. Hence, in order to account for the

noise present in the supply line, a guard-band is added on the core supply voltage so that the performance of the processor remains unaffected. This increase in the supply voltage, in turn, increases power dissipation by the cores.

In chapter 2 of the thesis, we study the sources of noises in the power supply network. We discuss prior arts with respect to reduction of noise from the supply network. We then propose an all-digital circuit technique based on staggering the activation of cores in a multi-core processor in order to reduce package inductor-induced power supply noise. In order to check the validity of the model, we implemented the concept in a test-chip built in 65nm bulk CMOS process, and measurement results show significant improvement in supply noise.

As mentioned earlier, for a processor with DVFS capability, as the number of cores per power domains reduces, power consumption reduces. However, this configuration requires more number of power domains. Greater number of power domains, in turn, requires larger number of supply lines. Supporting many supply lines from off-chip converters is prohibitive because of large off-chip component count, long response time of off-chip converters and large IR noise due to routing. These limitations can be overcome by building on-chip voltage regulators. Because of low quality of on-chip inductor, efficiency of an on-chip inductor-based switching regulator is usually very poor. Switched-capacitor based DC/DC converter, on the other hand, suffers from low power density. In our work, we use a novel 3-D capacitor, known as deep trench capacitor, to increase the power density of the on-chip switched capacitor DC/DC converter with reasonably high efficiency. Our design of DC/DC converters in chapter 3 and chapter 4 uses deep trench capacitor as the flying capacitors of the switched capacitor converter.

As discussed earlier, IR noise present in the supply line is one of the major performance limiting factors. In modern microprocessors, IR noise accounts for the 10% of the supply noise. Presence of large IR noise necessitates larger guard-banding on the supply voltage, thereby increasing power consumptions of the cores. Proper design of the package and the power supply grid can

result in reduction in the IR noise present in the supply line. However, we propose a circuit technique in order to further reduce the amount of IR noise present in the supply line. IR noise is proportional to the amount of current that is flowing to the cores from the voltage regulator through the power supply network. Hence, it is possible to reduce the amount of IR noise seen by the core by reducing the current flow though the power delivery network. It can be achieved by opportunistically borrowing a portion of the required current of the active core from an adjacent low-activity core. Since because of DVFS active core supply voltage is higher than the idle core supply voltage, we use an on-chip switched capacitor voltage doubler, which will step-up the voltage of the idle core to the voltage level of the active core before the idle core is able to supply current to the active core. We built and measured a test-chip in 32nm SOI process in order verify our proposed technique. The detail of this work can be found in chapter 3 of this thesis.

Another source of non-ideality in the supply line arises from the presence of large parasitic capacitance in the thick metal lines. DVFS requires scaling of supply voltage based on operating frequency. However, presence of large capacitance in the supply network slows down the charging and discharging of the supply network during DVFS transients. In order to solve the problem associated with slow charging of the supply network, we propose a switched capacitor based on-chip step-down converter which can be reconfigured with a boost mode in order to expedite the charging process of the supply network. This converter was built and measured 32 nm SOI process. The details of the converter can be found in chapter 3 of the thesis.

In chapter 4 of the thesis, we have shown the implementation details of a time division demultiplexed (TDM) single-input dual-output (SIDO) switched capacitor DC/DC converter built in 32nm SOI CMOS process. We have shown TDDM converter can achieve better power density than a conventional switched capacitor DC/DC converter under unbalanced load condition of its two outputs.

Recently, implementation of on-chip converters has been reported in the industry. Intel has built a fully integrated voltage regulator (FIVR) in its Haswell processor [29]. It is a switching voltage regulator that uses inductor of the package as the energy storage element. IBM has also reported the use of low-dropout (LDO) regulator based on-chip voltage regulation modules in its Power8 processor [30]. The use of on-chip voltage regulation schemes in the industry motivated us to compare various power delivery architectures in terms of total system power consumption. The details of this comparison work can be found in chapter 5 of the thesis. Under iso-throughput condition, we have compared average system power dissipation and minimum system power dissipation using FIVR and LDO based system architectures. We conclude that under the assumption of per-core DVFS, FIVR based power delivery architecture consumes less average power than LDO based one because of higher efficiency of FIVR. For the FIVR based architecture, if the number of cores per power domain increases to 2, this architecture still works better in terms of system power consumption as compared to a LDO based per-core DVFS architecture. We also conclude that FIVR based architecture is less susceptible to process variation than the LDO based one.

2. STAGGERED CORE ACTIVATION: A CIRCUIT/ARCHITECTURAL APPROACH FOR MITIGATING RESONANT SUPPLY NOISE ISSUES IN MULTI-CORE MULTI-POWER DOMAIN PROCESSORS

2.1 Overview and Motivation

Proper design of power delivery network is necessary to ensure low supply noise seen by microprocessor cores. Modern multi-core microprocessors, which operate at frequencies in the GHz range, are extremely susceptible to power supply noise. When supply voltage drops below its

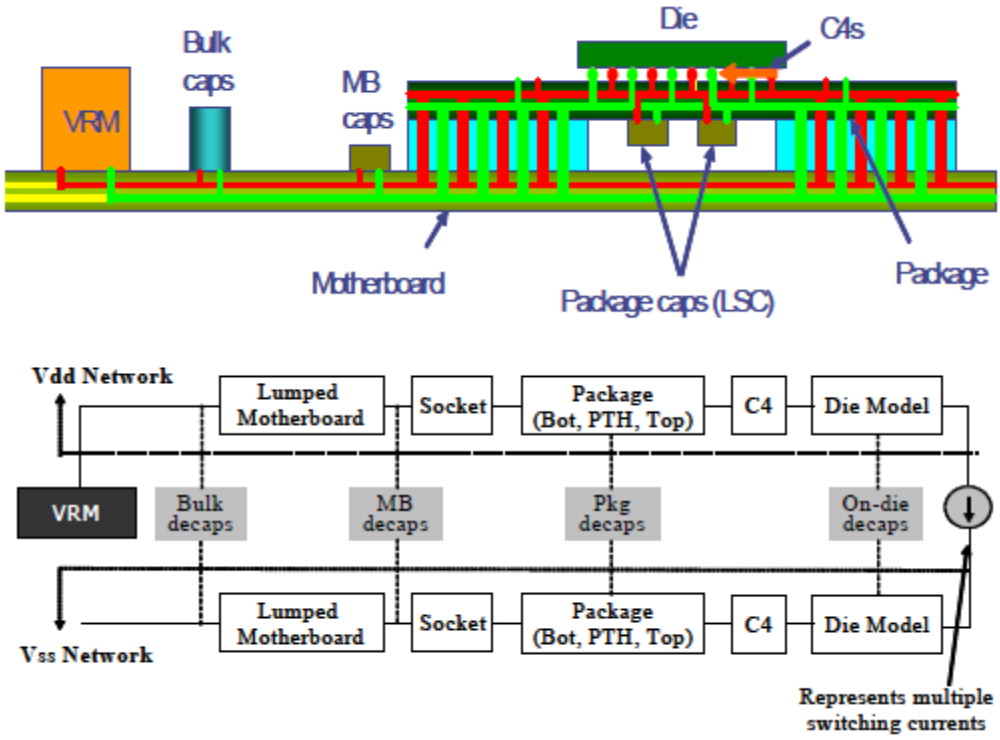


Figure 2.1: Typical power delivery network of microprocessor (Source: Intel) [45].

nominal value because of the presence of noise, digital circuits inside the cores may cease to operate correctly because of timing failures [5]. For proper functionality, affected circuits needs to be run at a lower frequency. Smaller operating frequency, in turn, affects the performance of the microprocessor negatively. On the other hand, supply voltage higher than the nominal value causes larger current drawn by the cores and creates reliability concerns like electromigration. Hence, the design of robust power supply network is one of major research areas in modern microprocessor designs [4].

When microprocessor cores draw current from power supply networks, supply voltage seen by the cores reduces by an amount IR because of the presence of parasitic resistance in the metals of the power supply network. Another important source of noise in the supply network is due to the inductive nature of the package wires. In this work, we concentrate our effort to reduce inductive noise from the power supply network.

Today's processors are equipped with various power-down modes. Clock gating reduces dynamic

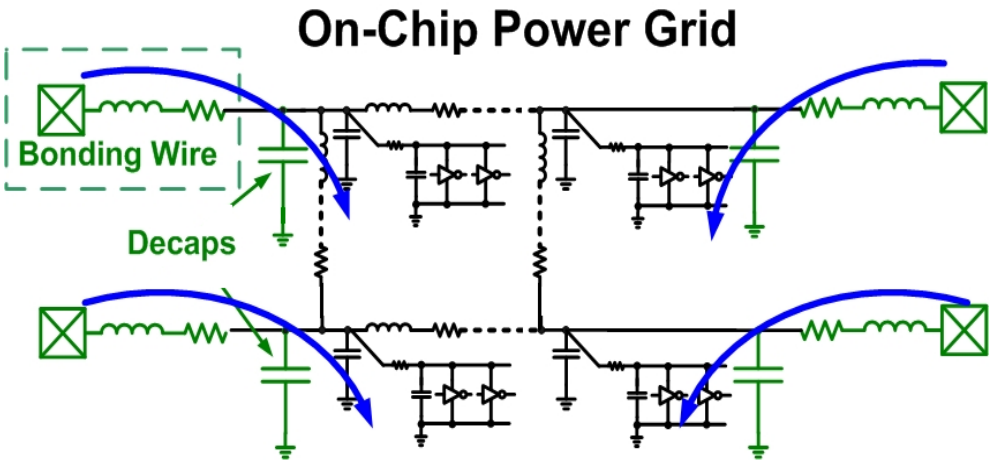


Figure 2.2: Formation of first-droop noise between package inductors and on-chip decoupling capacitors.

power consumption by turning the clock signals of the idle circuits off. On the other hand, power gating technique uses a sleep transistor which disconnects idle circuits from the supply network, thereby saving significant amount of leakage power. However, these low power techniques trigger a sudden increase or decrease in the current drawn from the supply line.

On-chip supply pads are connected to the package power grids by means of package wires in case of bond-wire technology, or by means of C4 (controlled collapsed chip connection) bump-array in case of flip-chip technology. These package leads are inductive in nature, and all of the currents drawn by the cores have to flow through these inductors. Hence, during core wake-up, core sleep, or transient clock events, package inductor experiences a sudden surge of current. The effect is manifested in a transient change in the supply voltage (V_{DD}), and the change in V_{DD} is given by the relation $\Delta V_{DD} = L\frac{di}{dt}$. Presence of parasitic capacitors in supply line and the non-switching devices in the circuit counteracts this transient voltage change. These capacitors are known as decoupling capacitors, or decaps. However, implicit device capacitance is not enough

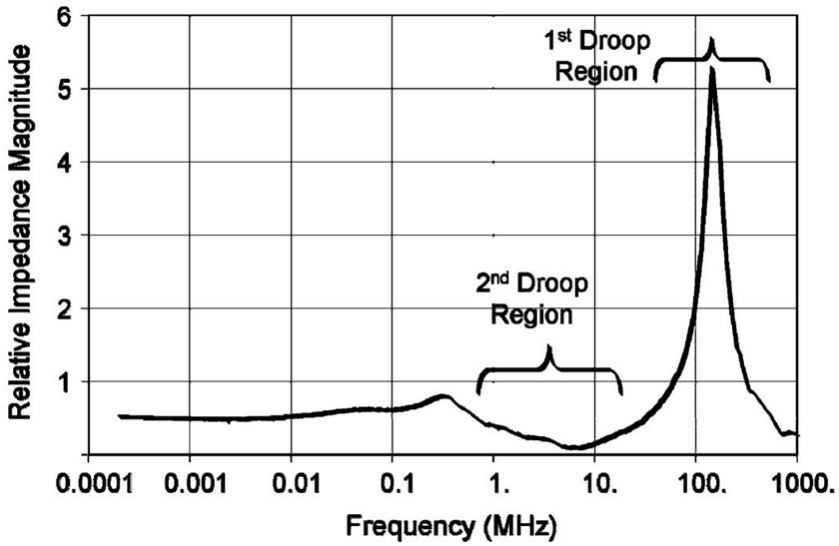


Figure 2.3: Measured power supply network impedance of Intel's Nehalem microprocessor. [9].

to ensure required supply integrity. Hence, additional decoupling capacitors needs to be added to suppress the effect of $L\frac{di}{dt}$. Since decaps take significant amount of on-chip area, they should be used sparingly.

Figure 2.1 shows power delivery network of a high performance microprocessor. Voltage regulator module (VRM), which sits on the motherboard, delivers power to the core through package wire or C4 bumps. In order to minimize $L\frac{di}{dt}$ noise, decaps are introduced at various locations, as can be seen from Figure 2.1. Resonant noise, which is caused due to the formation of L-C tank circuit between VRM and motherboard, and again between motherboard and microprocessor socket, is known as third-droop noise. L-C tank circuit between microprocessor socket and package is responsible for second-droop noise. First droop noise is formed between the package/bonding inductance and the die capacitance (see Figure 2.2), and it typically resides in the 40MHz to 300MHz frequency range. Finally, presence of significant inductance in the on-chip power grid causes a zeroth-droop noise.

Reduction of supply noise is one of the major concerns of IC design, and significant efforts have been made to mitigate the impact of supply noise on timing errors. A method to optimize placements of decoupling capacitors by using activity profiles has been proposed in [6]. In order to reduce resonant supply noise, an active damping circuit has been proposed in [7]. Gu proposed an active decoupling circuit to reduce the area of the decoupling capacitors and power consumption [8][10][12]. Circuit techniques like clock/data compensation method have been recently proposed to mitigate the effect of resonant supply noise [13]. However, only few attempts have been made to utilize the architectural level behavior of multi-core, multi-power domain processors which could be more effective, simpler, fully digital, and more scalable than previous brute-force circuit approaches.

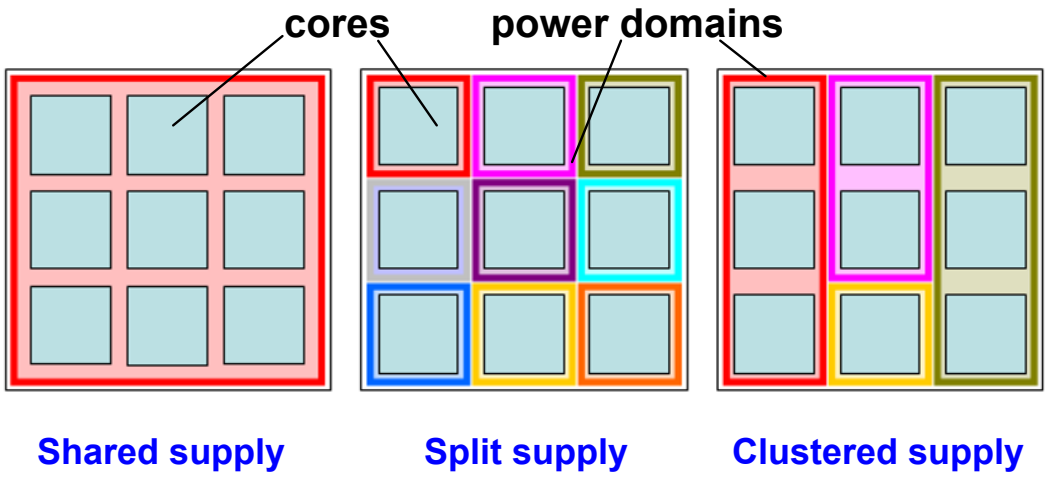


Figure 2.4: Multi-core power domain architectures.

In this chapter, we propose a circuit/architectural approach to reduce first-droop resonant supply noise in a multi-power domain processor by staggering the activation of the cores sharing the same power domain. Our work attempts to mitigate first-droop noise from an architectural perspective. Figure 2.3 shows the measured power supply network impedance of an Intel Nehalem microprocessor. It exhibits a large impedance peak around 150MHz [9]. [11] and [42] shows that resonant supply noise resides in the frequency range 40MHz-300MHz. Because of its global nature, this first droop noise affects the entire chip. It is large in magnitude, and hence, it constitutes the worst-case supply noise scenario.

2.2 Staggered Core Activation for Resonant Noise Reduction

Figure 2.4 shows three different configurations for multi-core processors. The leftmost configuration has all cores in a single voltage domain. This configuration is known as shared-supply architecture since all the cores in the processor share the same supply rail. As a result, individual control over core supply voltage is not possible with this configuration. Middle configuration of Figure 2.4 is known as split supply architecture. In this architecture, each core has its own supply rail. Although this configuration is useful for power saving since each core can run at its optimum supply voltage, because of area overhead and complexity in routing, split

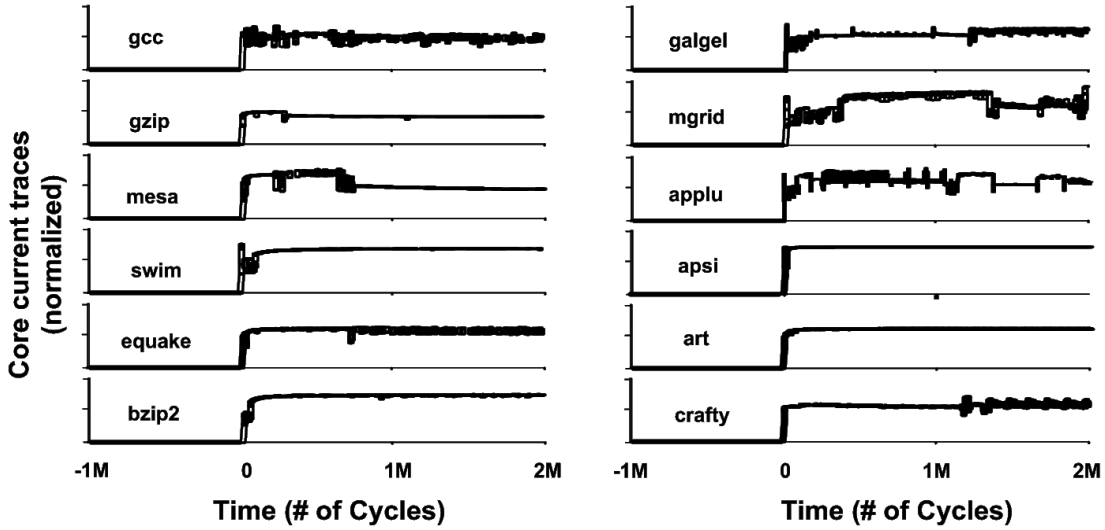


Figure 2.5: Normalized core current traces for various benchmarks.

supply architecture is not feasible with large number of cores in a microprocessor. Finally, the rightmost configuration in Figure 2.4 has a number of power domains, each of which is shared by a number of cores. This architecture is known as clustered supply architecture, since cores in the same cluster or power domain share the same supply voltage. Our approach is applicable within a single power domain of a multi-core multi-power domain processor, where all the cores share same power supply.

Typical workloads of a number of SPEC2000 benchmarks [14][15][16] are shown in Figure 2.5. All benchmark programs show a step function-like jump at the onset of running an application, and similarly an abrupt step down will occur upon their completion. Resonant noise excited by the step increase in the supply current at the instant of core wake-up creates damping sinusoid like ripple in the V_{DD} and ground network. The amplitude of the step depends on the amount of current drawn by a single core and on the number of cores going active at the same time. It is also dependent on the amount of on-chip decoupling capacitance and package inductance. It determines amplitude of the damping sinusoid noise.

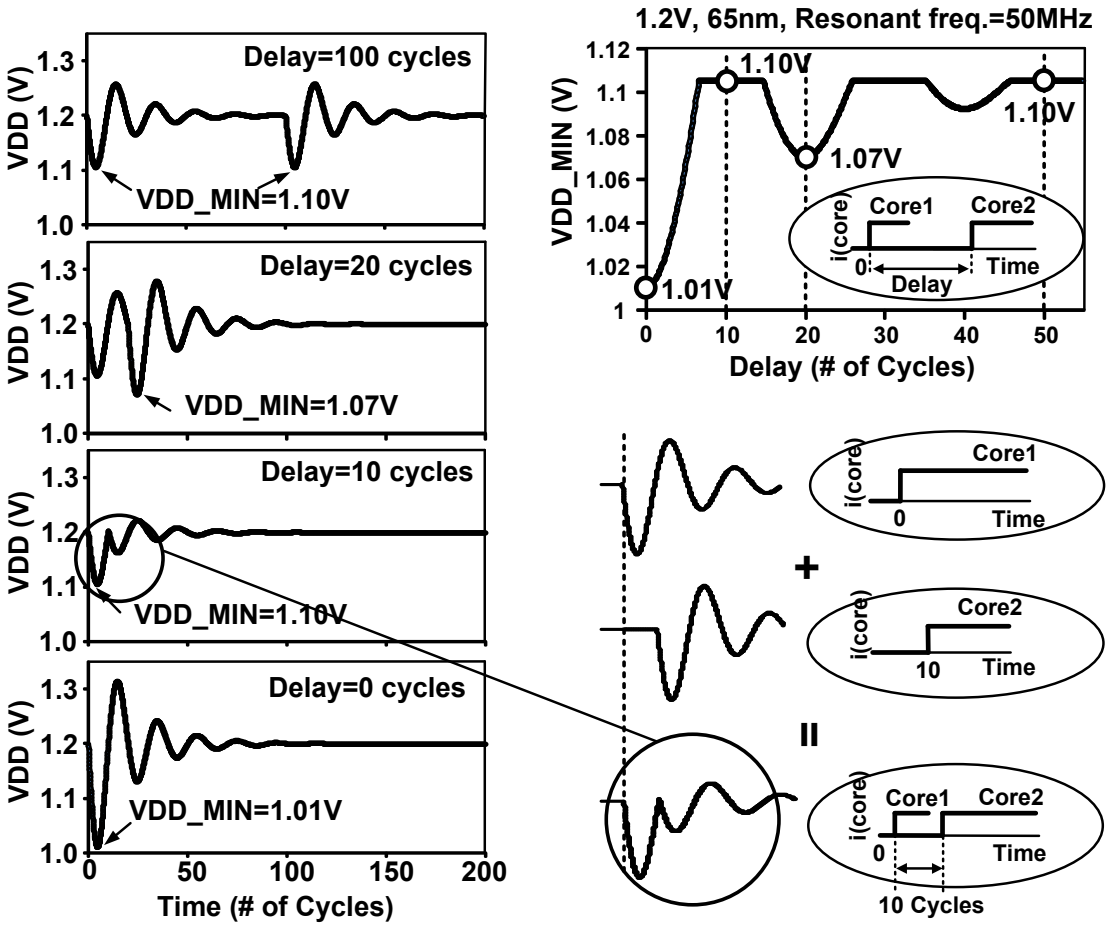


Figure 2.6: Graphical representation of stagger core activation scheme in a 2-core configuration.

Turn-on delay dependent resonant noise interactions of two cores are graphically represented in Figure 2.6. In case of a 2-core processor, if the turn-on times of the two cores are apart by several hundreds of cycle (Figure 2.6, top-left), the resultant noise is identical to the noise generated from a single core. However, as both the cores' turn-on times come closer, the noise waveforms start interacting with each other (Figure 2.6, mid-left), which reinforces or diminishes noise depending on the time difference between the turn-on times of the two cores. As one would expect, worst-case supply noise corresponds to the case when both the cores turn on simultaneously (Figure 2.6, bottom-left). As we sweep the turn-on delay of two cores and plot the minimum V_{DD} (V_{DD_MIN}),

we find a number of valleys and plateaus, corresponding to large noise and small noise, respectively. Our proposed Staggered Core Activation (SCA) scheme to reduce resonant noise is based on staggering the turn-on time of the latter core(s). A core requesting activation first will turn on immediately. However, latter core's activation request and its actual turn-on event may differ in time. If the latter core's activation request time corresponds to the valleys of the V_{DD_MIN} vs. delay waveform, we stagger the turn-on event of the latter core until the next plateau. This ensures minimum supply noise in a 2-core processor.

SCA can be extended to more than 2-core processor scenarios without significant modification. For an arbitrary N-core processor, core requesting activation first will turn on immediately. The core whose request comes second should wait until the plateau due to first core's turn on, as discussed in 2-core scenario. The core coming third in requesting activation should wait until the plateau due to second core's turn-on, and so on. Although for the absolute lowest possible supply noise, a core's stagger time may depend on all previous cores' turn-on times, its implementation will be too complicated while the effectiveness will diminish quickly.

2.3 Test Chip Implementation

A 65nm test chip was designed to verify the effectiveness of the proposed SCA scheme. The block diagram is shown in Figure 2.7. The test block consists of 16 simple cores that can be

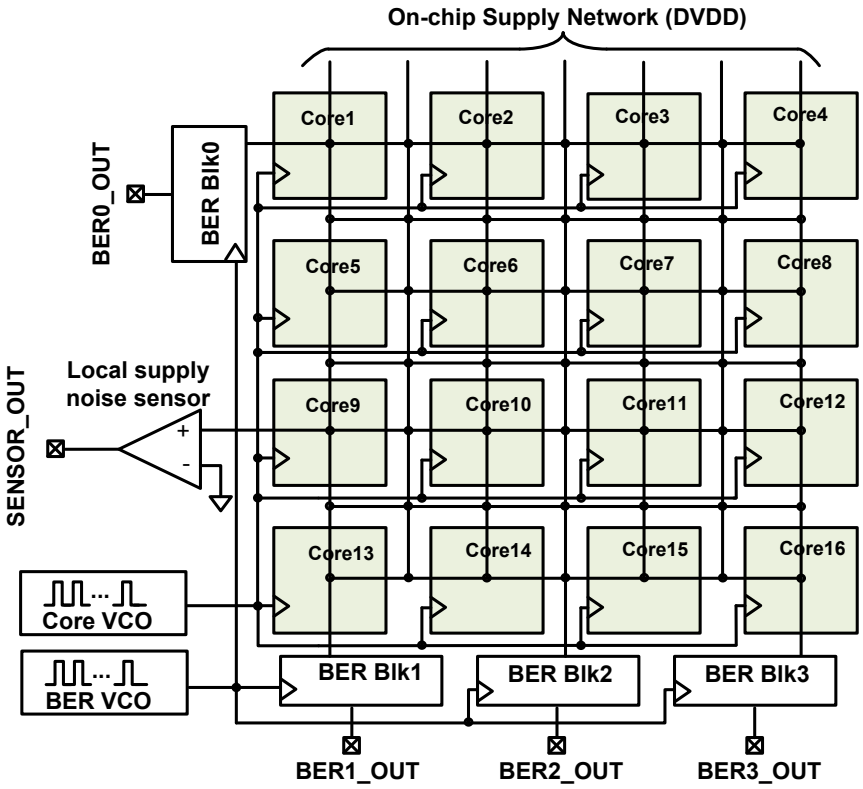


Figure 2.7: Block diagram of the 65nm test chip.

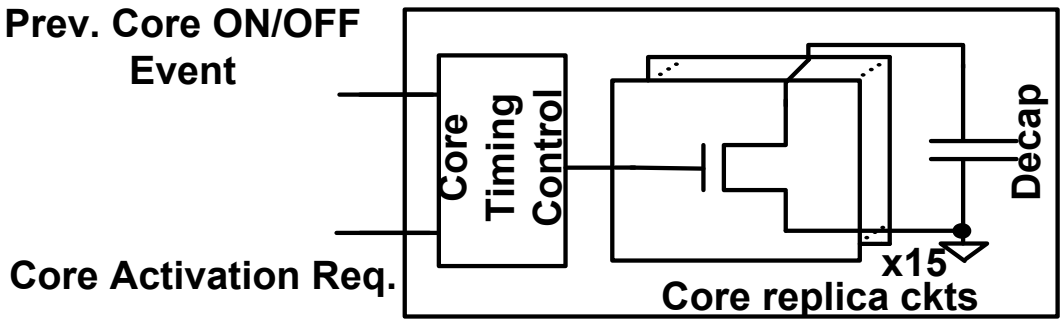


Figure 2.8: Replica of a Core.

activated at different times. Since the wake-up transition current of a core exhibits a step-like function, we use simple replica circuits such as individual devices and random logic blocks whose current profiles resemble that of typical processor cores as shown in Figure 2.5. The test block consists of 16 simple cores that can be activated at different times.

Each core has a precise timing control circuit and several noise injection blocks (Figure 2.8).

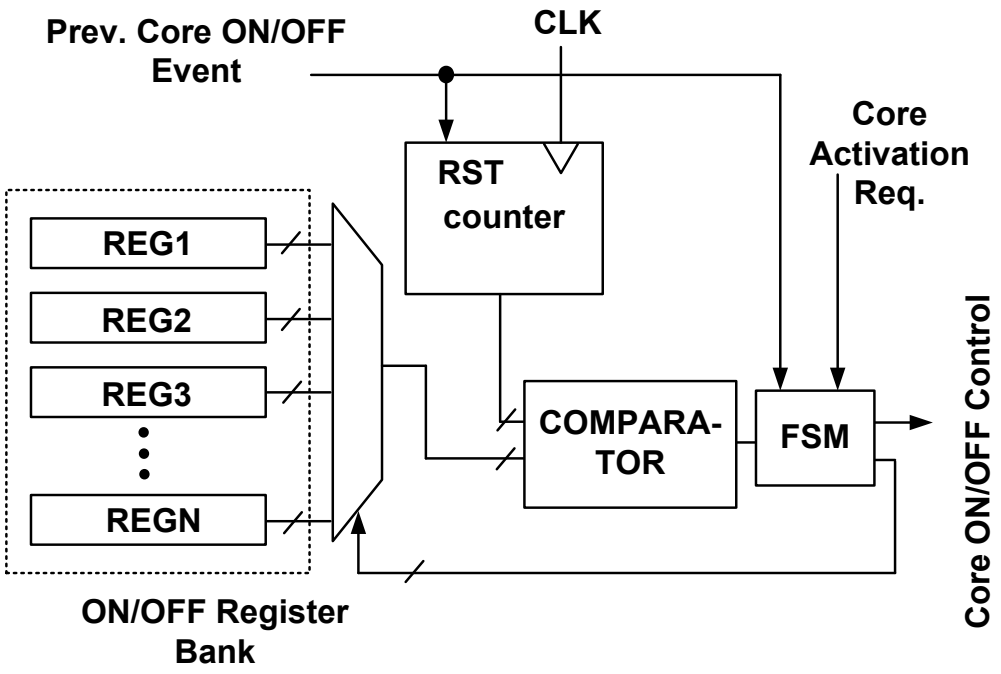


Figure 2.9: Core timing control unit for staggered core activation scheme.

Figure 2.9 shows a detailed block diagram of the timing control unit. The register bank inside the timing control unit stores the timing information corresponding to the beginning and the end points of the plateaus of the V_{DD_MIN} vs. delay waveform. Other core's ON/OFF event resets the counter of control unit and with current core's activation/deactivation request, counter output is compared with the proper register value and current core's ON/OFF control signal is sent out. A number of noise injection blocks inside the core give us the flexibility in terms of injecting noise of different amplitudes. These noise injection blocks are clocked by an on-chip VCO.

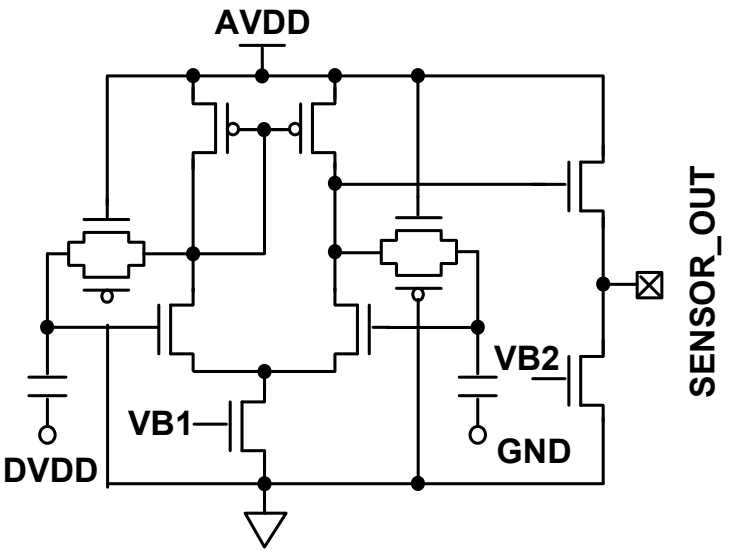


Figure 2.10: Schematic of a local supply noise sensor.

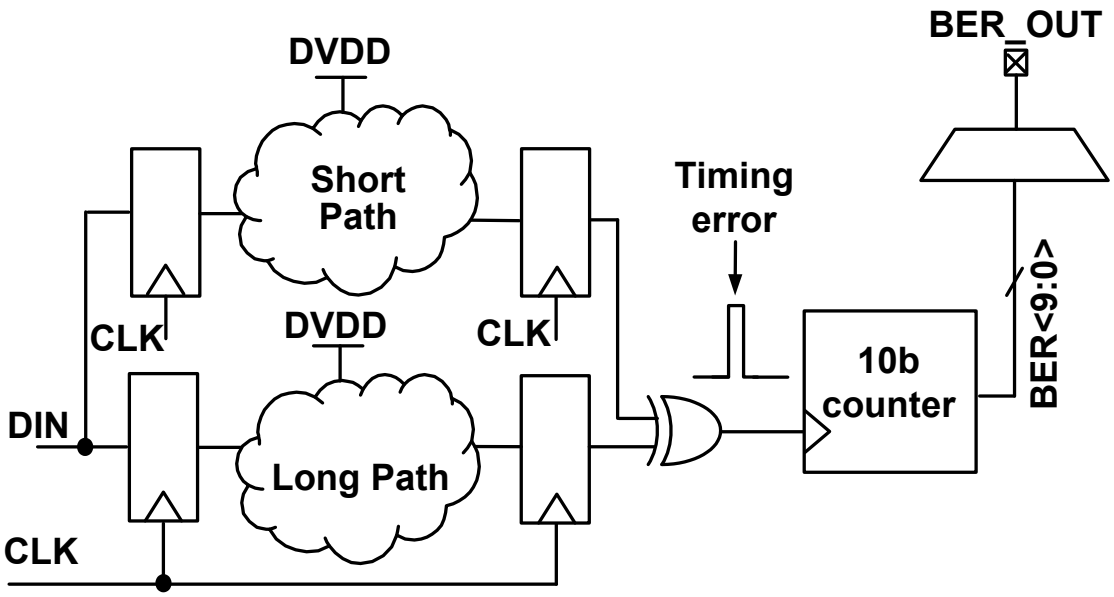


Figure 2.11: BER measurement block.

There is a local analog noise sensor for direct measurement of the local supply noise as shown in Figure 2.10. It takes the noisy supply and ground signals as the differential inputs, and its output indicates the on-chip supply noise frequency and amplitude [10]. The frequency response of the noise sensor can be characterized using external high frequency inputs and a spectrum analyzer.

To measure the Fmax improvement using SCA, four on-chip Bit Error Rate (BER) measurement blocks (Figure 2.11) with a dedicated VCO were added. The BER blocks count the number of errors produced as a result of supply noise by comparing the outputs of a very short logic and interconnect path and a long logic and interconnect path. By measuring the average period of the 10-bit ripple counter output and the VCO frequency, BER can be readily calculated. Changing the frequency that the BER block runs at will allow us to determine the Fmax for different time delay schemes. Here, without loss of generality, we assume Fmax to be the frequency at which BER=10^{-6}.

2.4 Test Chip Measurement Results

Figure 2.12 shows the oscilloscope capture of supply voltage for different delay numbers that we have obtained from our test-chip.

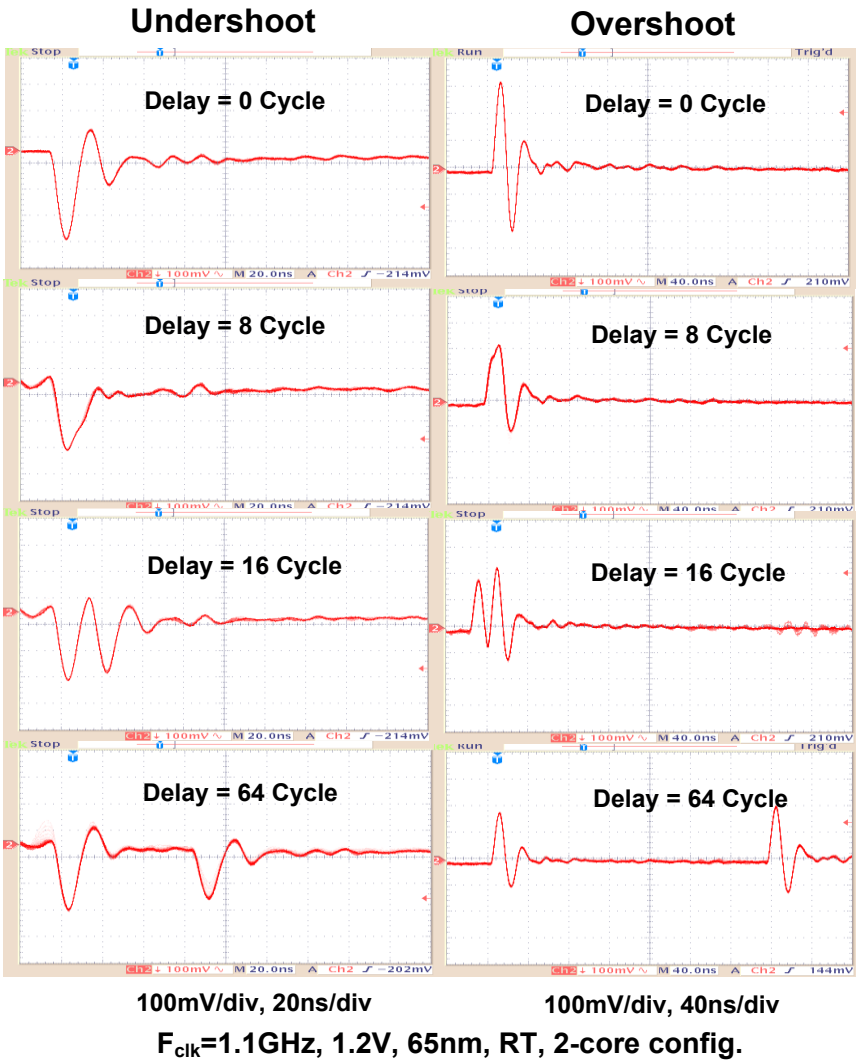


Figure 2.12: Example supply noise undershoots (left column) and overshoots (right column) as a function of delay in a 2-core configuration.

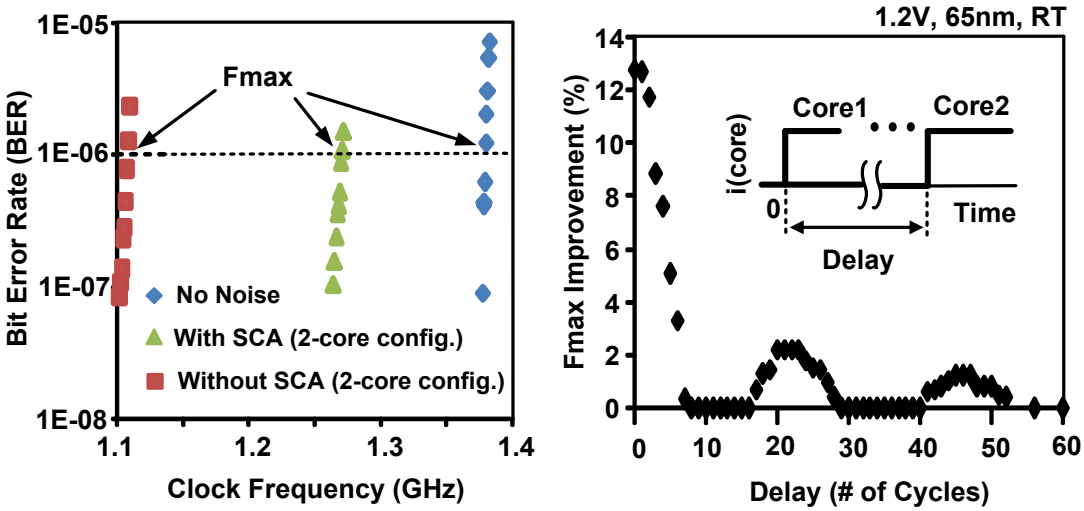


Figure 2.13: Measured BER versus clock frequency (left). Measured percentage Fmax improvement with SCA versus delay in a 2-core configuration (right).

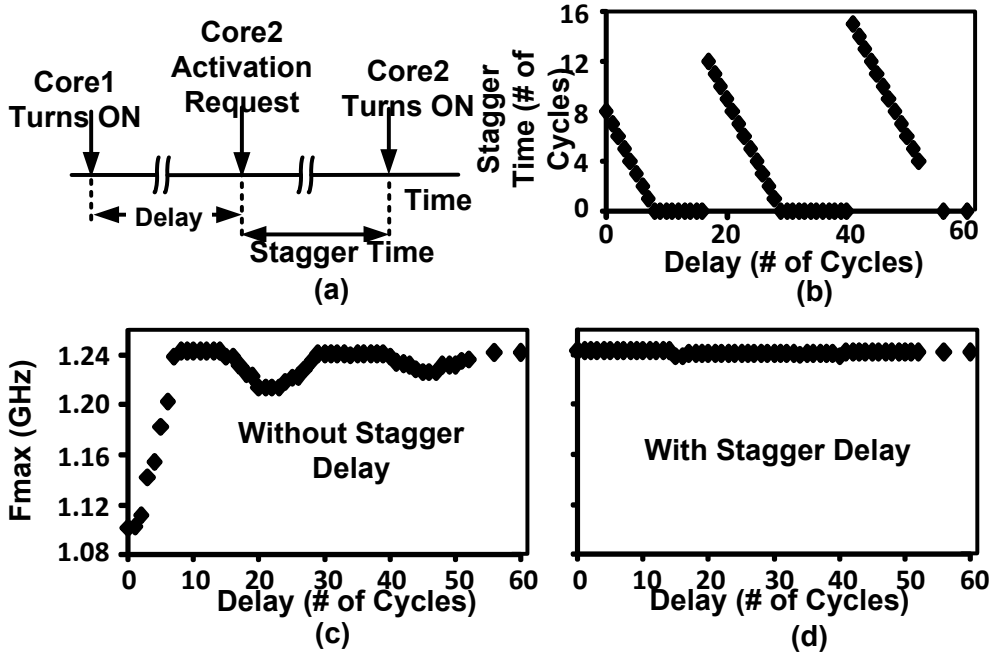


Figure 2.14: (a) Core activation time sequences in a 2-Core Processor. (b) Measured stagger time versus delay. Measured Fmax versus delay in a 2-core configuration (c) without SCA and (d) with SCA.

Figure 2.13 left plots BER vs. clock frequency for SCA, no SCA and no noise scenarios in 2-core

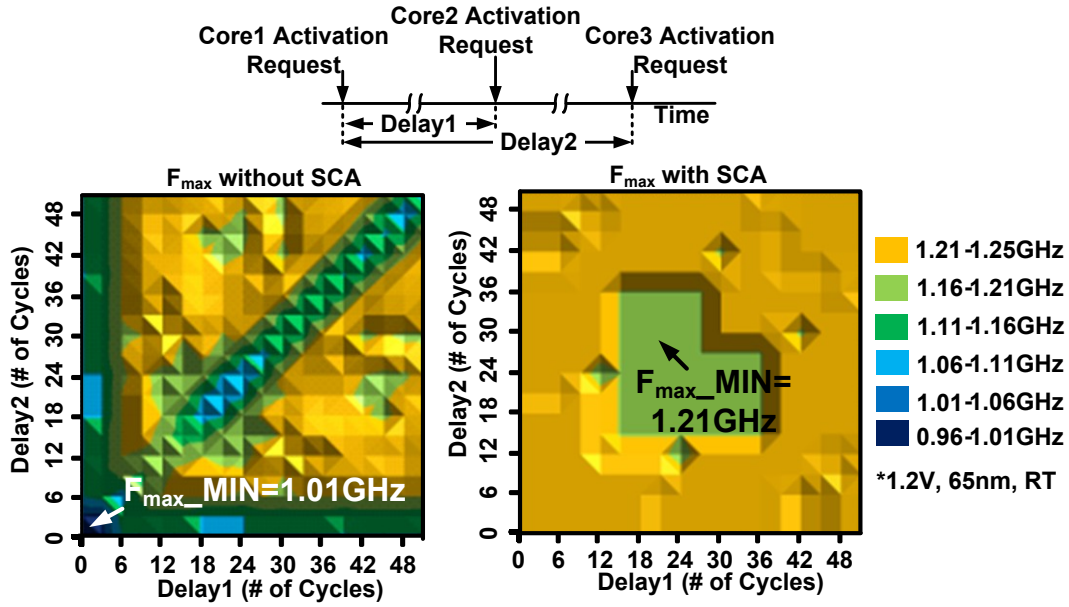


Figure 2.15: Core activation time sequences in a 3-core processor (top). Measured Fmax versus delay1 and delay2 in a 3-core configuration (bottom left) without SCA and (bottom right) with SCA.

processor. For a constant BER of 1E-6, Fmax could be increased to 1.37GHz for no noise case. However, it is ideal scenario, and in reality noise present in the supply line will determine the maximum clock frequency. From this figure, we see that for a constant BER value, proposed SCA core activation scheme results in larger clock frequency as compared to no SCA case. Figure 2.13 right shows the improvement in maximum clock frequency using SCA vs. turn-on delay. From this figure, we see that improvement is maximum when delay is 0 cycle. It is due to the fact that, for delay=0cycle, both cores turn-on simultaneously drawing huge amount of current from the supply line. Using SCA, we can avoid the worst-case noise, and get maximum improvement in Fmax. Also, we notice some regions of 0% improvement in Fmax in this figure. These regions correspond to the plateau regions of V_{DD_MIN} vs. delay waveform. It simply means that for these delay values, no further staggering is necessary as they already correspond to best-case delay values.

Figure 2.14 shows the measurement results of the SCA for a 2-core processor. Timing sequences for core turn-on events are shown in Figure 2.14(a). In Figure 2.14(b), the amount of stagger time given to the second core before it can turn on is plotted against delay between activation requests between the two cores. Without SCA, delay vs. Fmax shows certain valleys and plateaus (Figure 2.14(c)). Clearly, Fmax is worst (1.10GHz) when both cores turn on simultaneously. However, in the plateaus, Fmax increases to 1.24GHz. SCA staggers second core activation until the plateau regions starts, thereby ensuring 1.24GHz of Fmax independent of the activation request time of the second core (Figure 2.14(d)). Stagger time overhead for our measurements was found to be in

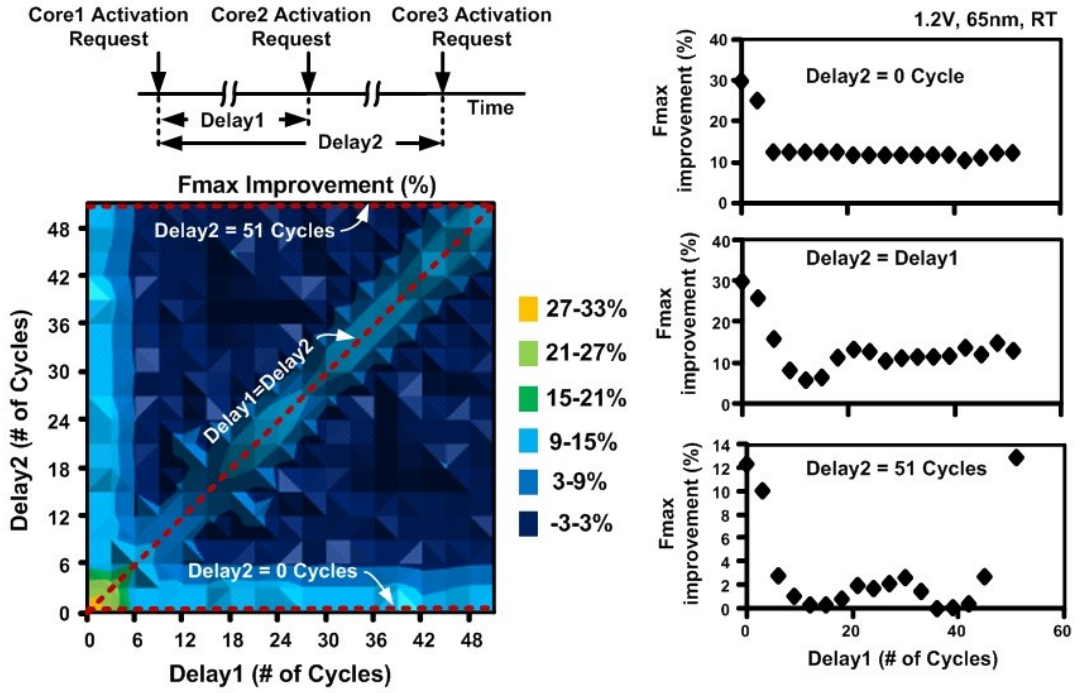


Figure 2.16: Measured percentage Fmax improvement with SCA versus delay1 and delay2 in a 3-core configuration (left). Measured percentage Fmax improvement with SCA versus delay1 in a 3-core configuration when delay2=0 (top right), delay2=delay1 (mid right), delay2=51 cycles (bottom).

the order of tens of cycles which is negligible compared to the typical number of cycles (>10^9) of a benchmark program.

Figure 2.15 shows measurement results of SCA in a 3-core processor configuration. Timing sequences for core turn-on events are illustrated in Figure 2.15 (top). Fmax improves from 1.01GHz when all three cores turn on at the same time (Figure 2.15 bottom left), to 1.21GHz with SCA (Figure bottom right). Comparison of the contour plots of with and without SCA indicates significant increase in Fmax in the former for different activation request time of core2 and core3. For a processor with more number of cores (N>3), Fmax without SCA is worse because of even larger current during simultaneous all-core activation, and with SCA we get more percentage improvement in Fmax.

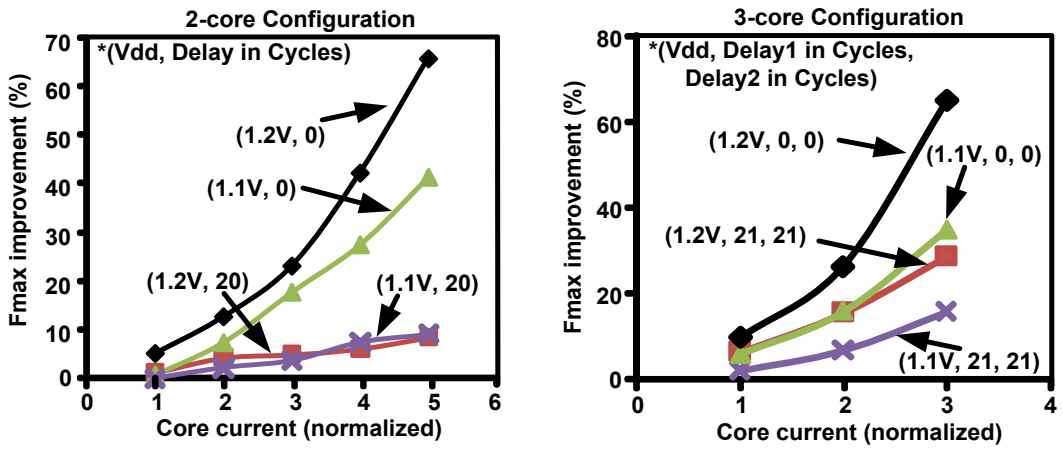


Figure 2.17: Measured percentage Fmax improvement with SCA versus normalized core current as a function of V_{DD}.

Plots in Figure 2.16 represent percentage improvement in a 3-core configuration as functions of delay1 and delay2. In the right hand side, we have plotted percentage improvement in Fmax for 3-different delay scenarios, e.g., delay2=0, delay2=delay1 and delay2=51 cycles.

Delay2=delay1=0 cycles corresponds to the worst-case noise scenario, and hence this case shows maximum improvement.

Figure 2.17 shows Fmax improvement with core current as a function of V_{DD}, for a 2-core and a 3-core configuration. For a fixed V_{DD}, percentage improvement in Fmax increases with increasing current as one would expect. For a fixed core current, as V_{DD} goes low, core frequency reduces, and percentage improvement in Fmax reduces.

Test chip microphotograph is shown in Figure 2.18.

Thus, in this work, we have presented a novel circuit/architectural approach based on staggering activation of cores in a multi-core multi-power domain processor that reduces the resonant supply noise. A 1.2V, 65nm test chip demonstrates approximately 12% and 20% performance improvement in Fmax for a 2-core processor and a 3-core processor, respectively. We have also shown that Fmax improvement using the proposed staggered core activation scheme is larger as the amplitude of the current drawn from the supply increases. This approach can be further extended to the case of many-core processors, where we expect to see a further improvement in terms of Fmax.

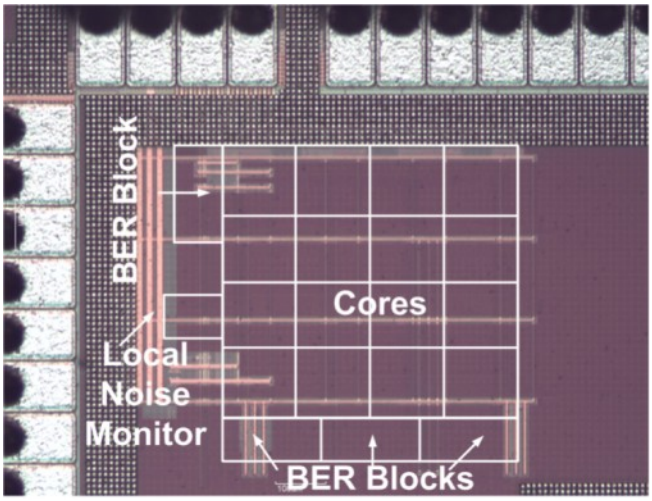


Figure 2.18: Test chip microphotograph.

3. DEEP TRENCH CAPACITOR BASED STEP-UP AND STEP-DOWN DC/DC CONVERTERS IN 32NM SOI WITH OPPORTUNISTIC CURRENT BORROWING AND FAST DVFS CAPABILITIES

3.1 Introduction and Overview of Switched Capacitor DC/DC Converter

DVFS has become one of the key low-power techniques in modern microprocessors. However, in order to exploit the full benefit of DVFS, it is necessary to be able to control the supply voltage of each core independently. With large number of cores on-chip, it became infeasible to build off-chip voltage converters for each core because of large off-chip component count. On the contrary, on-chip voltage converters have the advantage of low component count, low IR drop, and better transient characteristics. Because of these advantages, there has been a flurry of researches on on-chip voltage regulators over the past few years [17][23][24][25][29][30]. In this section, we study the basics of switched capacitor based on-chip DC/DC converter. Although [43] has reported an on-chip inductor based switching DC/DC converter, the efficiency of the converter suffers because the quality of on-chip inductor is poor. Linear regulators are simple to implement on-chip. However, their efficiency degrades as the output voltage of the converter deviates from its input voltage. Switched capacitor DC/DC converters solve the problems associated with both types of converters. However, since capacitor charging-discharging process is inherently lossy, they cannot achieve very high efficiency like inductor based converters. Also, because of low on-chip capacitor density of MOS capacitors and MIM (Metal-Insulator-Metal) capacitors, high power density is not achievable from them.

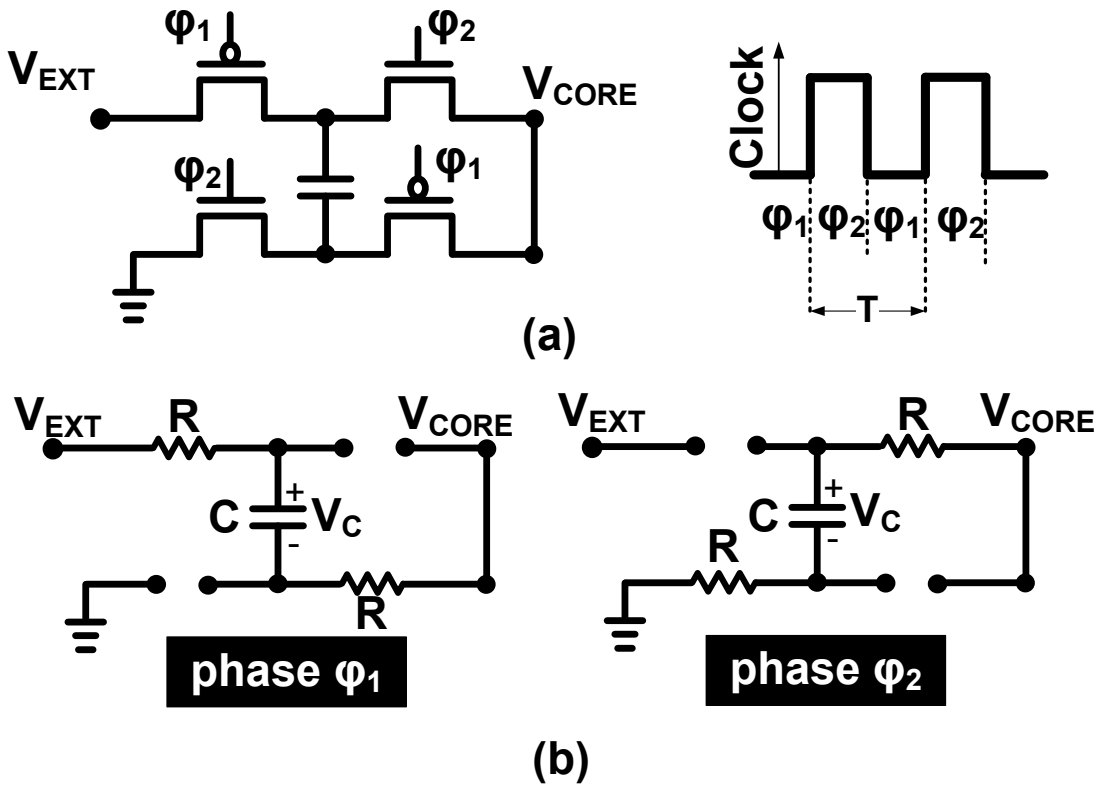


Figure 3.1: (a) Schematic of 1:2 switched capacitor DC/DC converter with clock signals shown on top right. (b) Switch configurations of the converters for both clock phases.

Figure 3.1(a) left shows the schematic of a 1:2 converter. Here, V_{EXT} is the input voltage of the converter, and V_{CORE} is the voltage generated by the converter. PMOS switches are driven by φ_1 phase of clock, whose period is T. During this phase, two NMOS switches remain off. Assuming R is the ON resistance of the switches, equivalent RC circuit for the converter in phase φ_1 is shown in Figure 1(b) left. During φ_2 phase of clock cycle, NMOS switches remain ON, whereas both PMOS switches are OFF. Corresponding RC circuit is shown in Figure 1(b) right. We assume that each of these phases is equal in time.

When the converter is in steady-state, we assume that voltages across the capacitor C are V_{C1} and V_{C2} at the end of phase φ_1 and φ_2, respectively. Hence,

$$V_{C1} = (V_{EXT} - V_{CORE}) + (V_{C2} - (V_{EXT} - V_{CORE})) \cdot e^{-\frac{T}{4RC}}$$

$$V_{C2} = V_{CORE} + (V_{C1} - V_{CORE}) \cdot e^{-\frac{T}{4RC}}$$

Average current that can be obtained from the converter can be given by the following relation:

$$i_{AVG} = \frac{1}{T}\int_0^T C\frac{dV_c}{dt} \approx C \cdot \frac{2}{T} \cdot (V_{C1} - V_{C2})$$

Since, $V_{C1} - V_{C2} = (V_{EXT} - 2 \cdot V_{CORE}) \cdot \frac{1-e^{-\frac{T}{4RC}}}{1+e^{-\frac{T}{4RC}}}$, average current can be given by $i_{AVG} = 2 \cdot f_{sw} \cdot (V_{EXT} - 2 \cdot V_{CORE}) \cdot \frac{1-e^{-\frac{T}{4RC}}}{1+e^{-\frac{T}{4RC}}}$, where $f_{sw} = \frac{1}{T}$ is known as the switching frequency of the converter.

DVFS has become a popular approach to improve the performance of microprocessors, especially for multi-core multi-power domain processors, while keeping an acceptable power consumption budget. However, long voltage switching time and large supply noise are major performance limiting factors of per-core or per-cluster DVFS. In this work, we propose a circuit technique based on switched capacitor DC/DC step-down converter to improve DVFS response time significantly. We also propose to use switched capacitor DC/DC step-up converter for supply noise reduction in a multi-power domain scenario by implementing a bi-directional opportunistic current borrowing scheme. Deep trench capacitors, which can be associated with "More-than-Moore" paradigm of ITRS roadmap, are used in our proposed DC/DC converters as flying capacitors. Originally meant for high density embedded memory applications, deep trench capacitors are slowly finding their ways into real systems, e.g. they are being used in DC/DC converters [17], PLL loop filters [18], decoupling circuitry [19] etc. Because of their inherent 3-D nature, large capacitance can be realized using a significantly smaller silicon footprint as compared to a metal-oxide-semiconductor (MOS)/ metal-insulator-metal (MIM) capacitors. They

are more than 20X denser than MOS capacitors, and hence can show significant improvement in power density when used in DC/DC converters [17].

3.2 Deep Trench Capacitor as Flying Capacitors

Figure 3.2 shows a cross-sectional view of a deep trench capacitor along with a 4x4 deep-trench array layout with trenches and contacts highlighted. Sidewalls of silicon (Si) trenches are arsenic (As) implanted to form the bottom plate of the capacitor, on top of which high-k dielectric materials are deposited. Finally, the trenches are refilled with As-doped polysilicon, which form

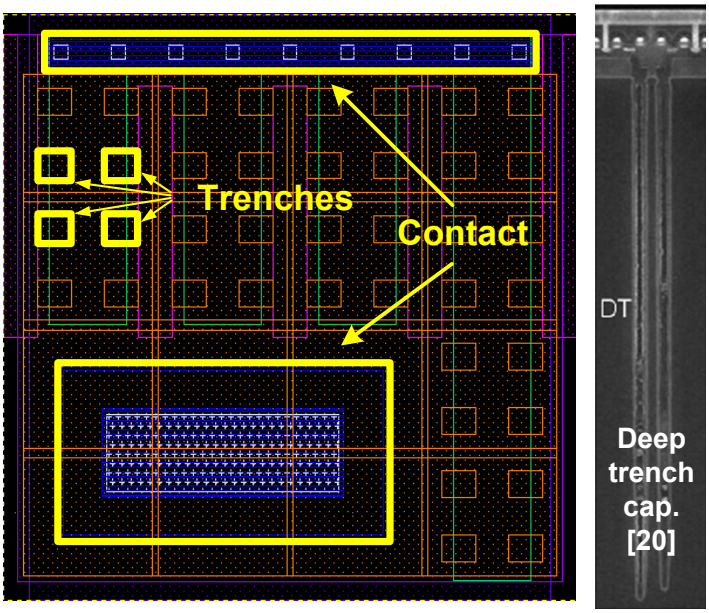


Figure 3.2: Layout and cross-sectional view of a deep trench capacitor.

the top electrode of the capacitor [21]. Large capacitor density of a deep trench capacitor comes with its large inherent series resistance, which is quite big in comparison to the equivalent series resistance (ESR) of MOS/MIM capacitors. [21] presents a distributed RC model in order to characterize trench capacitors.

Trench fill resistance and plate resistance of the outer electrode are the sources of parasitic resistance of deep trench capacitors. To a first order approximation, a deep trench capacitor can

be modeled as an ideal capacitance in series with an ESR, whose frequency responses are shown in Figure 3.3(a). Because of the distributed nature of deep trench capacitors, ESR and C_{ideal} are

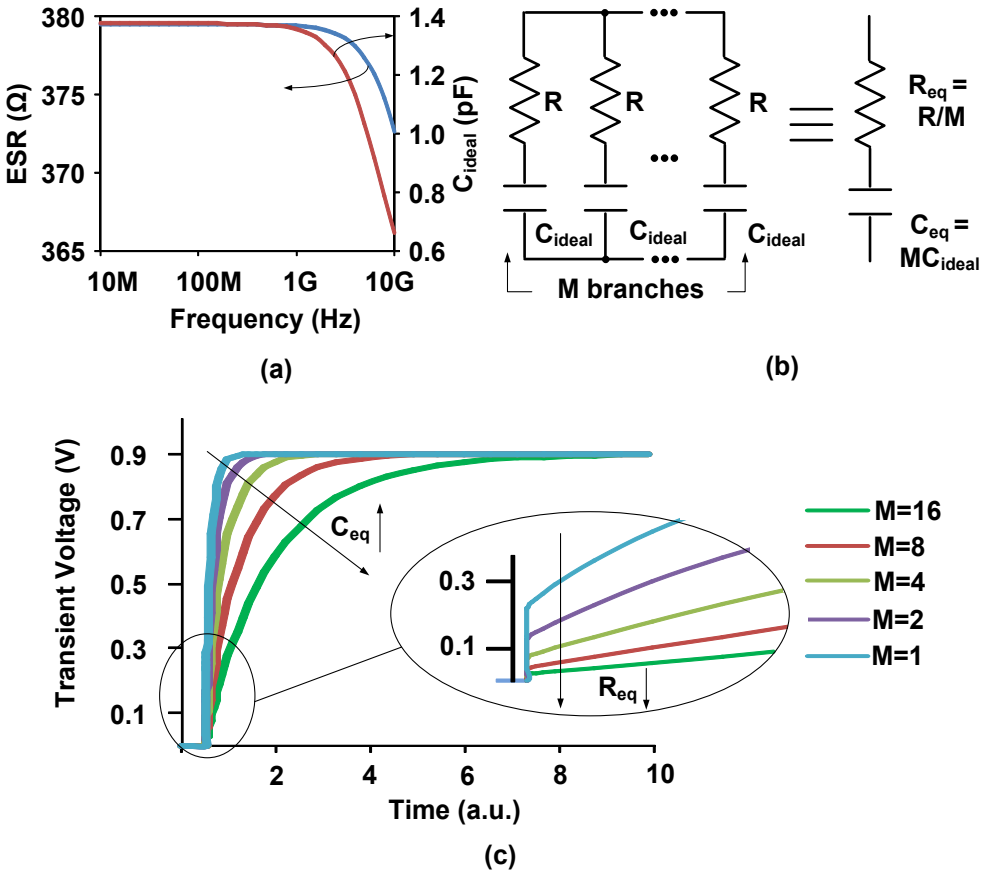


Figure 3.3: (a) Frequency response of a deep trench capacitor (6x6 array). (b) ESR reduction with multiplicity. (c) Transient response of a deep trench capacitor as a function of capacitor multiplicity.

found to decrease at ultra high frequency, although they are practically frequency independent within our operating frequency range. However, because of the presence of large ESR, a deep-trench capacitor exhibits slower transients than an equivalent MOS capacitor. Also, the presence of high ESR affects the efficiency of deep trench capacitor based DC/DC converters adversely. However, in order to supply even a light load (~10μW), necessary flying capacitors can be

realized by connecting multiple deep trench arrays in parallel, which in turn lowers the effective series resistance of the structure by a factor of multiplicity of the arrays as shown in Figure 3.3 (b). Figure 3.3 (c) shows transient response of a deep trench capacitor as a function of the multiplicity (M) of the capacitor. Although our design is virtually insensitive to the high ESR associated with a deep trench capacitor array, designers should carefully consider the impact of high series resistance on circuit performance before using deep trench capacitors.

3.3 Fast DVFS Scheme Using Boost Mode

Block diagram of proposed step-down converter has been shown in Figure 3.4. 32 converter modules operate in a time-interleaved manner in order to ensure smaller output ripple. The voltage request sent by the microprocessor (V_{ref}) is compared with the output voltage of the converter (V_{out}). Output of the comparator provides proper bias to a voltage-controlled oscillator (VCO) to generate multi-phase complementary clock signals which drive the step-down

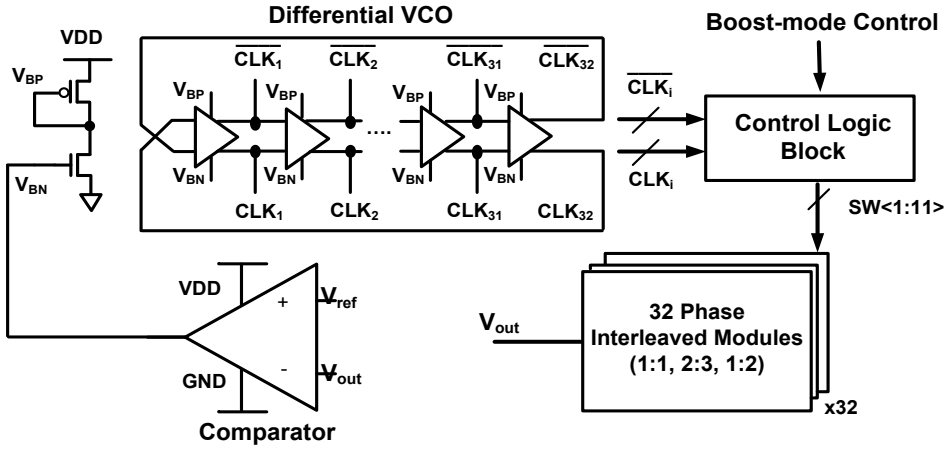
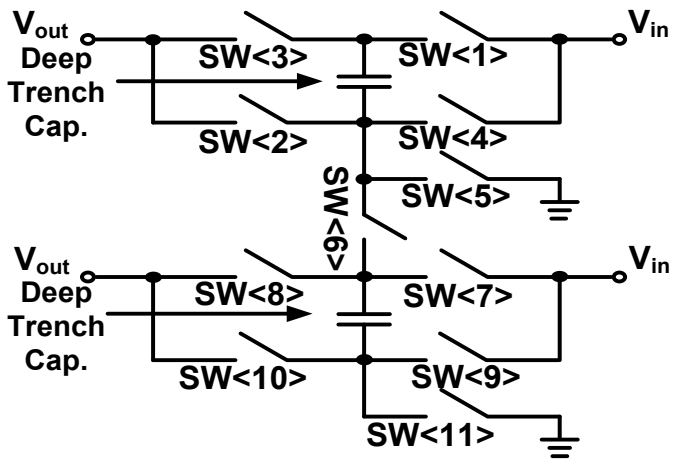


Figure 3.4: Block diagram of step-down converter in a feedback control loop.

converter. The number of stages of the VCO is selected as large as possible to achieve better multi-phase interleaving for the step-down converter block. On the other hand, it should also satisfy the requirement of the maximum operating frequency, which is determined by the trade-off between power density and efficiency. The power consumption of the VCO needs to be



SW	1:2		2:3		Boost (1:1)	
	Phase1	Phase2	Phase1	Phase2	Phase1	Phase2
1,2,7,10	Close	Open	Close	Open	Close	Open
3	Open	Close	Open	Close	Open	Close
4,9	Open	Open	Open	Open	Open	Close
5	Open	Close	Open	Open	Open	Open
6	Open	Open	Open	Close	Open	Open
8	Open	Close	Open	Open	Open	Close
11	Open	Close	Open	Close	Open	Open

Figure 3.5: Basic step-down converter module showing switch configurations for different conversion ratios.

minimized to optimize the overall efficiency of the proposed converter. Proposed reconfigurable step-down converter using deep trench capacitor is capable of generating V_{out} with conversion ratios 1:1, 2:3 and 1:2. The usefulness to generate three voltage ratios from a single converter is that, it can cover a greater range of output voltage without sacrificing efficiency significantly. Schematic of a step-down converter module along with its switch configurations to generate above-mentioned output ratios are shown in Figure 3.5.

Our fast DVFS scheme using step-down converter has been explained in the flowchart shown in Figure 3.6. Sluggish response due to voltage level rise during DVFS can be attributed to the

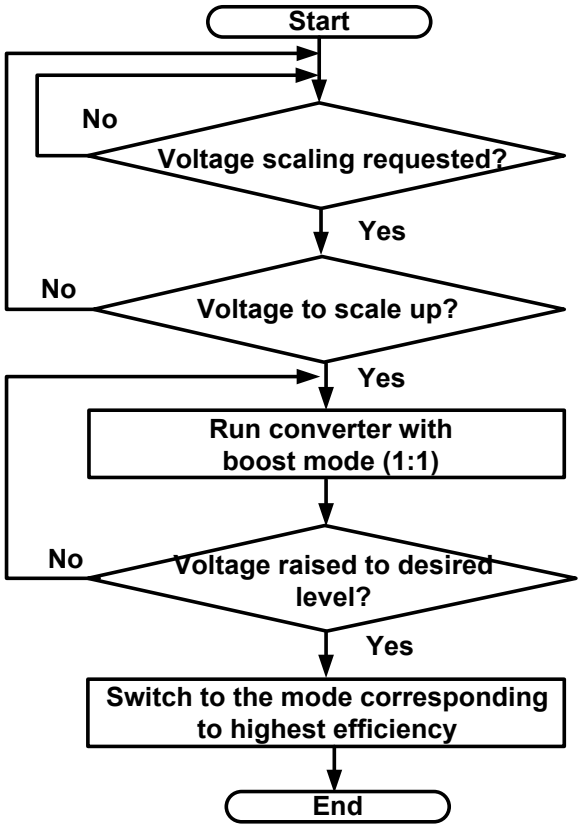


Figure 3.6: Proposed fast DVFS control scheme with boost mode operation.

charging time of the large output load, which can be greatly reduced by pumping more charge during voltage rise transient. With each request to raise the output voltage level, control logic of the step-down converter configures itself to 1:1 ratio. Since the amount of charge delivered per unit time using 1:1 configuration is maximum compared to the charge delivered using other smaller ratios, time required to charge output load reduces. However, once the converter reaches desired voltage level, switches are reconfigured according to the closest higher output configuration in order to ensure high efficiency.

3.4 Supply Noise Reduction by Opportunistic Current Borrowing

Proposed supply noise reduction technique works on the basis of borrowing current from adjacent idle/lightly loaded cores and dumping it to an active core. Although borrowed current will lead to an extra IR drop on adjacent cores, performance degradation in those cores will be negligible

since those cores are running under light-load condition. One thing to note here is that the supply

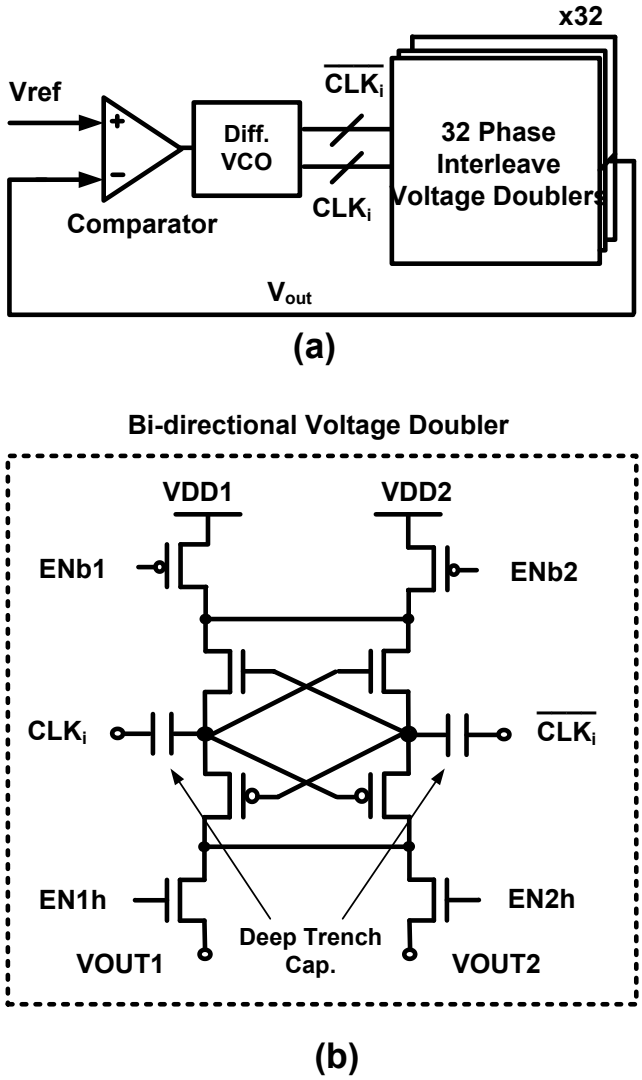


Figure 3.7: (a) Step-up converter block diagram with feedback and (b) Modified Favrat cell for bi-directional voltage doubling.

voltage of the adjacent idle cores can be lower than that of the active core due to the nature of DVFS. Therefore, the voltage levels of the idle cores must be boosted before they can provide current to the active core. Based on this observation, we propose to use a step-up converter, which uses deep trench capacitors as flying capacitors, to achieve this goal. Figure 3.7 (a) shows

the step-up converter consisting of voltage doubling units, a differential VCO for generating multiphase clock and a comparator connected in a feedback loop. Modified Favrat cells (Figure

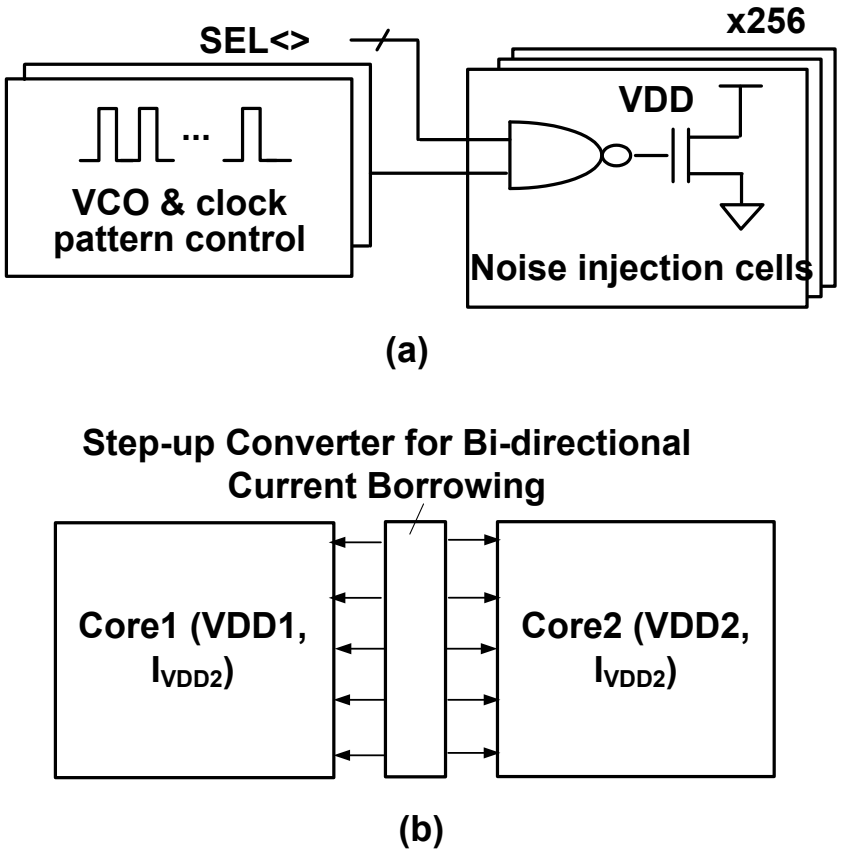


Figure 3.8: (a) On-chip supply noise generation circuits for test purposes. (b) Per-core DVFS architecture with proposed opportunistic current borrowing step-up converter.

3.7 (b)) are used for bi-directional voltage doubling [22].

As shown in Figure 3.8 (a), core replica circuits consist of a number of noise injection blocks whose turn-on times can be precisely controlled. These noise injection blocks are clocked by an on-chip VCO, and they are able to generate noise of different patterns and amplitudes. Two different power domains (VDD1, I_{VDD1} and VDD2, I_{VDD2}) (Figure 3.8 (b)) interact with each other with the help of the voltage doubler placed between them. By controlling switches (EN1h, EN2h,

ENb1, ENb2) of the bi-directional voltage doubler unit cell shown in Figure 3.7 (b), it can work in three different modes: (1) VDD1 provides current to boost VDD2; (2) VDD2 provides current to boost VDD1; (3) and a disabled mode. This flexibility in supplying current through multiple

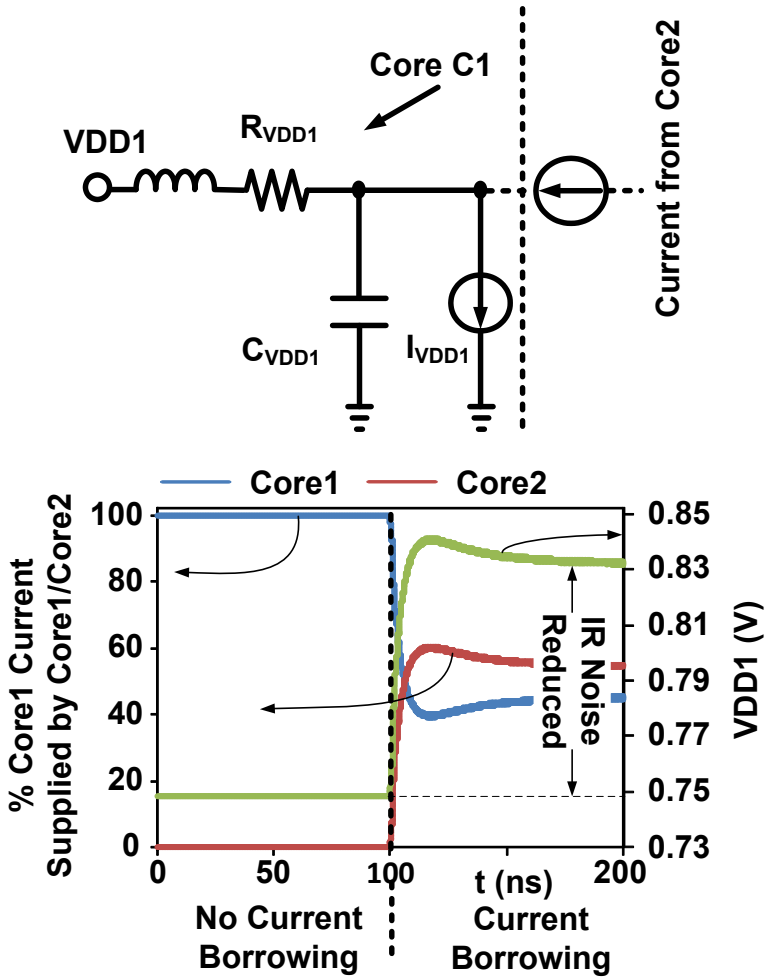


Figure 3.9: Equivalent circuit and simulation waveforms explaining proposed current borrowing scheme.

power domains provides an additional control knob which helps improve the overall power efficiency of the processor. Equivalent circuit and simulation waveforms showing IR noise reduction of core1 by borrowing current from core2 have been presented in Figure 3.9.

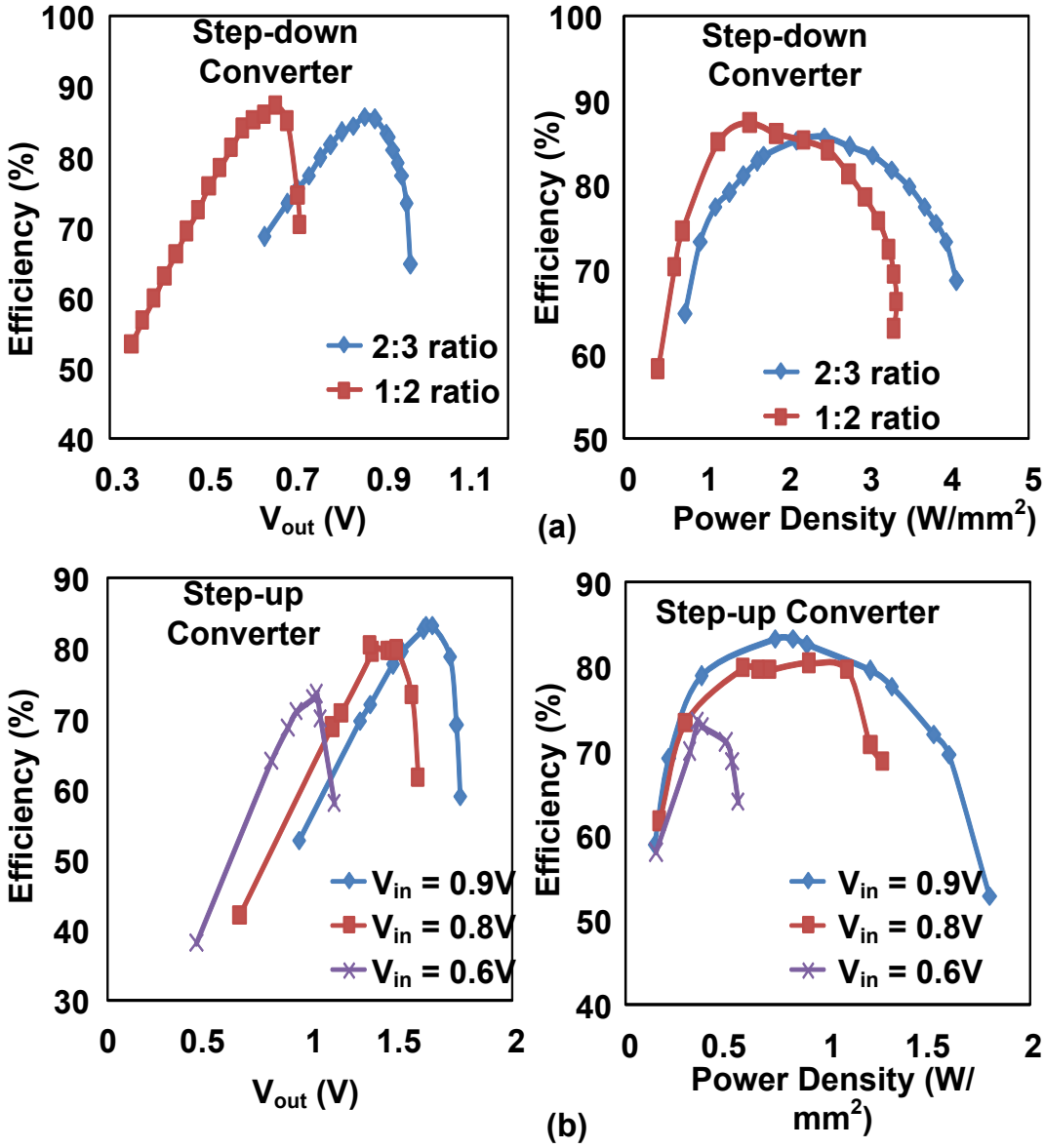


Figure 3.10: Open-loop efficiency vs. Vout and open-loop efficiency vs. power density for (a) step-down converter at 110MHz and (b) step-up converter at 24MHz.

3.5 Deep Trench Capacitor Based Converter Measurements

Open-loop efficiency vs. V_{out} and open-loop efficiency vs. power density plots of step-down and step-up converters are shown in Figure 3.10(a) and Figure 3.10(b), respectively. Efficiency plots

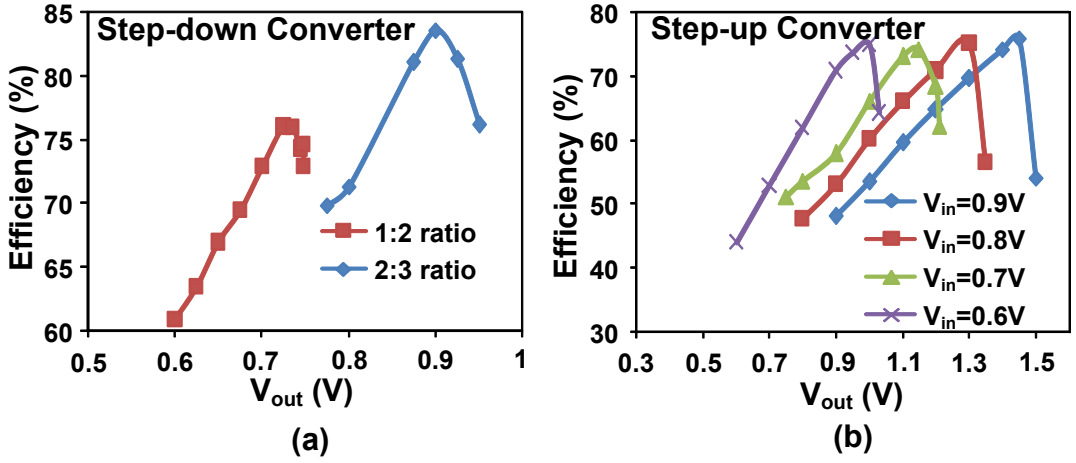


Figure 3.11: Closed-loop efficiency vs. V_{out} of (a) step-down and (b) step-up converter.

for the step-down and the step-up converters are obtained at optimum frequencies of 110MHz and 24MHz, respectively. Efficiency vs. V_{out} plots for various input voltage levels (V_{in}) of the step-up converter show largest efficiency with smallest V_{in} for V_{out} in the range between 0.7V and 1.0V. This implies that for a nominal supply voltage of 0.9V of the active core, the smaller the supply

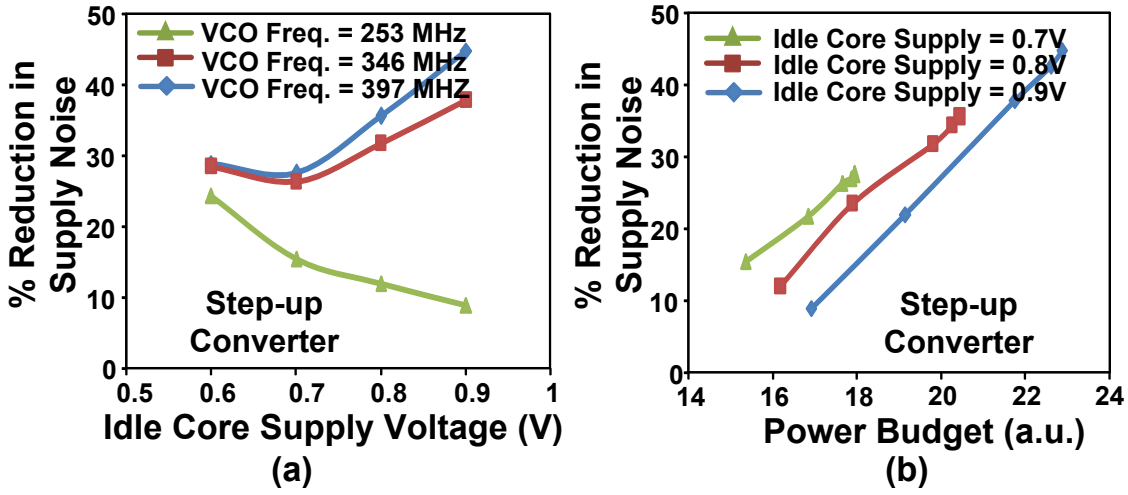


Figure 3.12: (a) Percentage reduction in supply noise vs. idle core supply voltage and (b) Percentage reduction in supply noise vs. power budget.

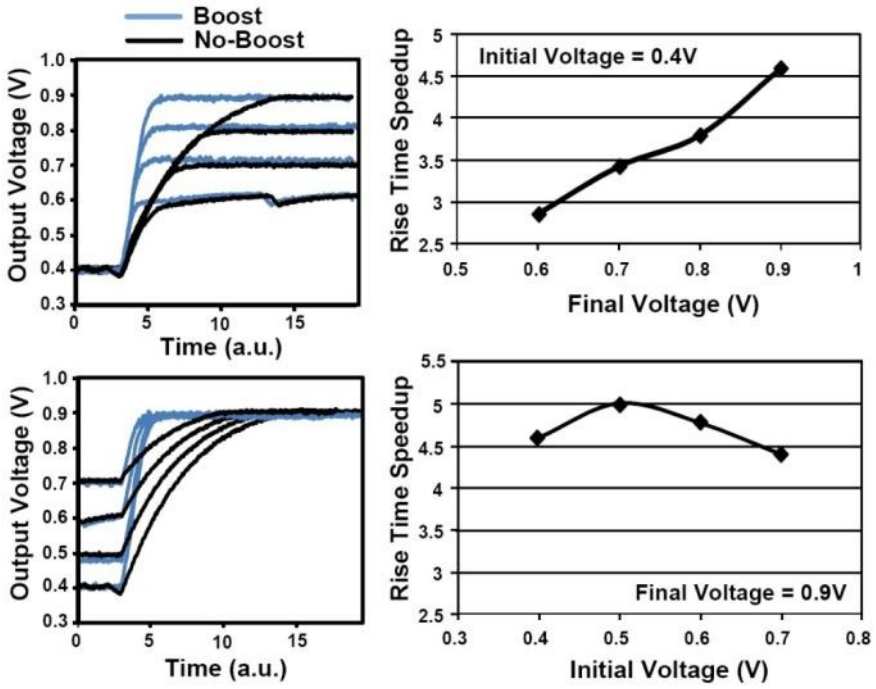


Figure 3.13: Measured waveforms showing rise transients of step-down converter with and without boost mode (left). Rise time speedup as functions of final Vout and initial Vout (right).

voltage of the idle core is, the better the efficiency of the current borrowing scheme will be. With an input voltage of 0.9V, the step-up converter is able to deliver a power density of 0.9W/mm^2 at a peak efficiency of 82% while maintaining efficiency over 70% throughout the voltage range from 1.25V to 1.73V. Step-down converter, on the other hand, uses an input voltage of 1.5V, and is capable of delivering 2.78W/mm^2 at a peak efficiency of 85% when it is configured as 2:3. It has efficiency over 70% in the voltage range from 0.5V to 0.98V. Closed-loop efficiency vs. V_{out} plots of the step-up and step-down converters assuming ring oscillator loads are shown in Figure 3.11. Closed-loop efficiency numbers take the loss in the feedback control loop into account, and hence are smaller compared to their open-loop counterpart.

In order to evaluate proposed current borrowing scheme for IR noise reduction, we have plotted percentage improvement in noise as a function of supply voltage levels of idle core for various

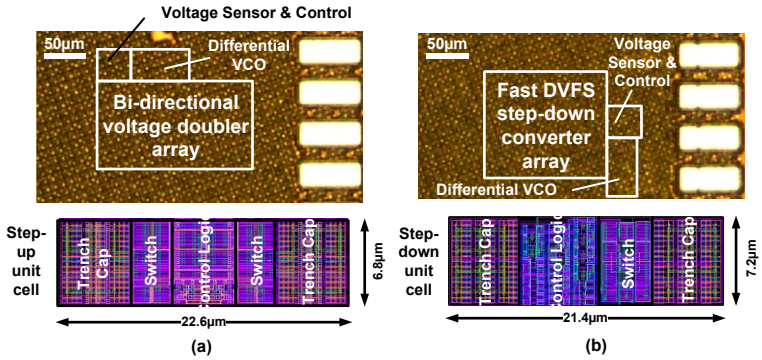


Figure 3.14: Test chip microphotographs of (a) step-up and (b) step-down converters in 32nm SOI with unit converter cell layout.

VCO frequencies of the step-up converter (Figure 3.12 (a)), and as a function of power budget for various supply voltage levels of the idle core (Figure 3.12 (b)). Larger VCO frequency leads to faster charge transfer to the active core, whereas limited power budget poses a current limit on the idle core, thereby limiting percentage noise improvement.

Measurement results for fast DVFS scheme have been shown by the voltage transients for constant initial and different final voltage levels (top left of Figure 3.13), constant final and different initial voltage levels (bottom left of Figure 3.13). Ratio of switching time with no-boost scheme and with boost scheme has been plotted as a function of varying final voltage (top right of Figure 3.13) and varying initial voltage (bottom right of Figure 3.13). We get 5X improvement in switching time for the case when V_{out} rises from 0.5V to 0.9V.

Test chip microphotographs highlighting major blocks of the step-up and step-down converters are shown in Figure 3.14 alongside the unit cell layouts. A summary of our step-up and step-down converters and their performance in comparison to the current state-of-the art switched capacitor converters are tabulated in Table 3.1.

Thus, in this work, we have presented switched capacitor based step-down and step-up converters using deep trench capacitors as flying capacitors. Simulation results have verified that with larger multiplicity of deep trench capacitors, effective ESR reduces by the same factor, and its effect on switching transient of the converter becomes negligible. We have designed the step-down converter for high speed DVFS by operating the converter with highest possible ratio (1:1 in our case) during rise transient of output voltage. On the other hand, step-up converter has been implemented for IR noise reduction of a core by borrowing current opportunistically from adjacent low activity cores. Measured data from a 32-nm test chip have shown a power density of 0.9W/mm^2 at a peak efficiency of 82% from the proposed step-up converter, and a power density of 2.78W/mm^2 at a peak efficiency of 85% from the proposed step-down converter. We have measured 45% reduction in IR noise using opportunistic current borrowing scheme, which can further be improved with larger VCO frequencies, although the power budget will act as a bottleneck at higher frequencies. From the test chip, a 5X improvement in rise transient has also been measured when V_{out} of the step-down converter is made to change from 0.5V to 0.9V.

Table 3.1: Comparisons with recent switched capacitor converter designs

	Tech. Node	Flying Cap.	Operating Modes	Feedback Control Loop	Max. Eff.	Power Density at Max. Eff.	Additio-nal Circuit Feature
[17]	45nm SOI	Deep Trench Cap.	2:1 Step-up & 1:2 Step-down	No	90%	2.3W/mm^2	None
[23]	32nm SOI	MOS Cap.	2:3, 1:2, 1:3 Step-down	No	81%	0.55W/mm^2	None
[24]	32nm Bulk	Metal Finger Cap.	2:1 Step-up	No	60%	1.12W/mm^2	None
This work	32nm SOI	Deep Trench Cap.	2:1 Step-up & 1:2, 2:3, 1:1 Step-down	Yes	85%	2.78W/mm^2 for 2:3 ratio	Fast DVFS, IR noise reduction

4. A TIME DIVISION DEMULTIPLEXED SINGLE-INPUT DUAL-OUTPUT DISTRIBUTED SWITCHED-CAPACITOR DC-DC CONVERTER UTILIZING DEEP TRENCH CAPACITOR IN 32NM SOI

4.1 Introduction and Overview

Over the past few years, on-chip switched capacitor (SC) DC/DC converter [17][23][25] has been researched as one of the low IR noise alternatives to traditional off-chip inductor based DC/DC buck converter because of the close proximity of the former to the processor core. The use of novel on-chip capacitors such as deep trench capacitors [17] and ferroelectric capacitors [25] has been reported, which can substitute traditional MOS/MIM capacitors as the flying capacitors. Research has also been directed towards implementing various SC DC/DC converter architectures for efficiency improvement, ripple reduction and power density improvement. Although single input multiple output (SIMO) DC/DC converter built using a single inductor has been studied thoroughly [26][27], to the best of our knowledge, no work has been reported so far on SIMO architecture for SC DC/DC converter. In this work, we propose a deep-trench capacitor based single-input dual-output (SIDO) SC DC/DC converter which uses time division demultiplexing (TDM) in order to provide power to its two outputs.

The architecture consists of two cores, which are powered up by switched capacitor converters. It is shown in Figure 4.1 top. For this purpose, conventional architecture uses two separate converters, each one acting as a supply for one core (Figure 4.1 top left). Here, supply voltage and load current for core1 are V_{CORE1} and I_{CORE1}, respectively. Corresponding values for core2 are V_{CORE2} and I_{CORE2}. Maximum power that could be delivered to each output is limited by sizes of the capacitors of the converters and the switching frequency. For a constant output power, if switching frequency is increased, the amount of

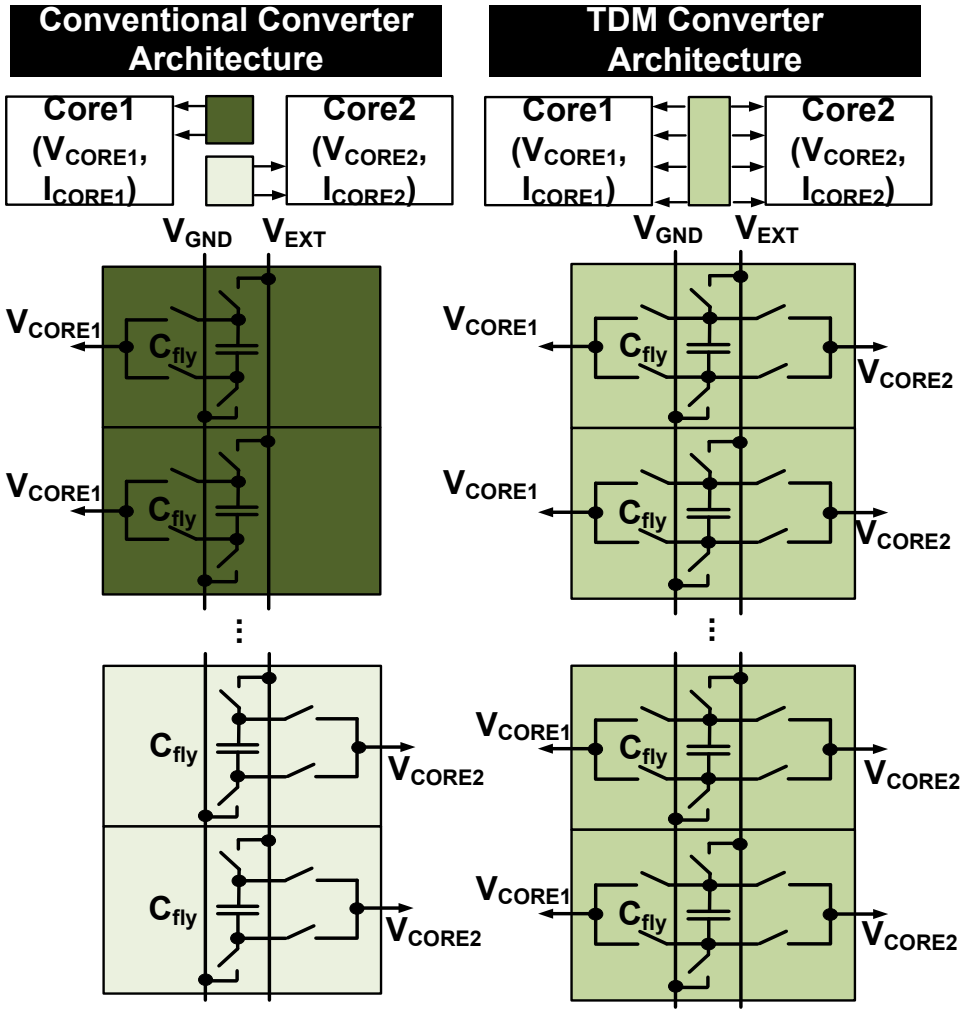


Figure 4.1: Architecture and switch schematic of a conventional converter and a TDDM converter.

required capacitor can be reduced by the same factor. However, this relation is true only in the slow switching regime, i.e. when the time period of the clock signal that drives the switches in the converter is larger than the RC time constant of the converter, where C is the flying capacitor of the converter and R is the total parasitic switch and wire resistances of the converter. Hence, in this slow switching regime, by increasing the frequency, it is possible to dump more charge to the output. However, in the fast switching regime, RC time constant is larger than the time period of the clock signal. Hence, flying capacitor does not get enough time to discharge enough before it starts charging again. As a matter of fact, in the fast switching regime, voltage across capacitor does not change significantly, and hence, the

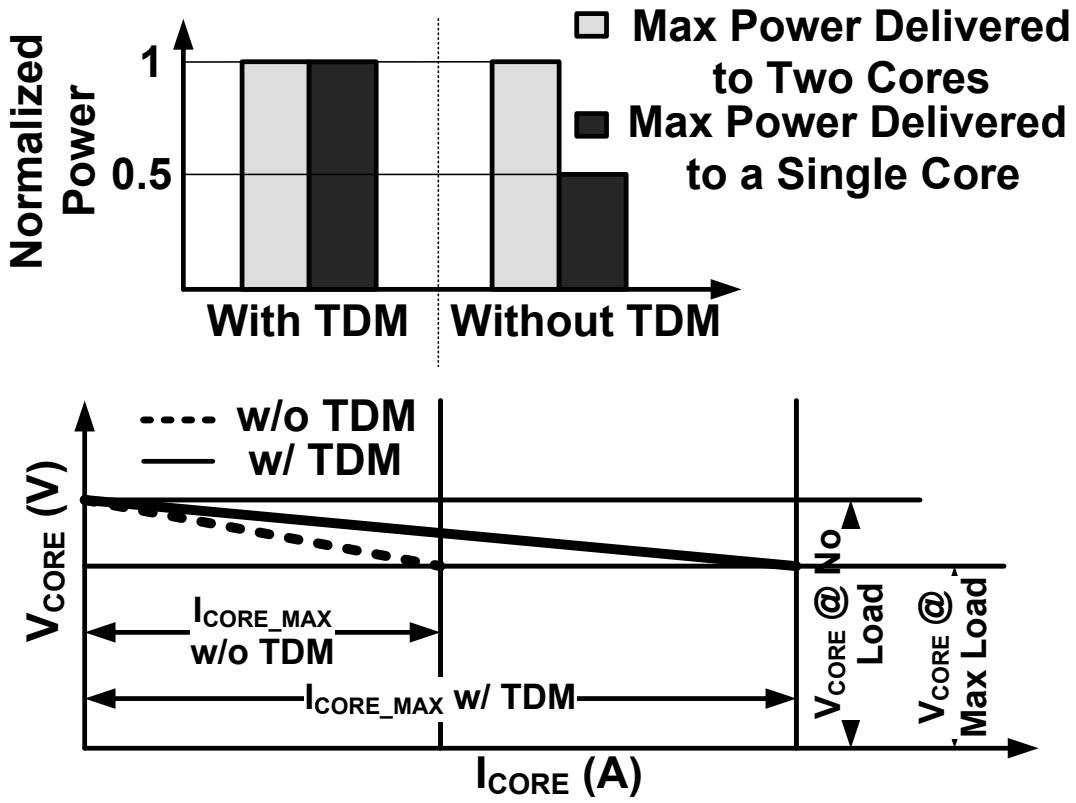


Figure 4.2: Comparison of load currents from TDDM converter and a conventional converter.

capacitor acts as a constant voltage source. Because of this behavior, output power delivered by the converter cannot be increased indefinitely by increasing the frequency.

For a switching converter, one major loss component is the loss occurred during driving the NMOS and PMOS switches. If the gate capacitance of the switch transistor is C_{GATE}, gate driving loss of each switch can be given by $C_{GATE}V_{DD}^2f_{sw}$, where V_{DD} is the amplitude of the clock signal driving the gate, and f_{sw} is the switching frequency of the converter. From the relation, we see that the larger switching frequency causes larger switching loss due to capacitor charging-discharging phenomenon through the parasitic resistances of the switches, and deteriorates efficiency of the converter. Hence, for a converter, which is limited by the efficiency, output power cannot be increased by increasing switching frequency. Output

power of the converter is also dependent on the amount of flying capacitance. Larger flying capacitance ensures larger output power at the cost of more silicon area.

Our proposed TDDM architecture provides power to multiple outputs from a single converter, as shown in Figure 4.1. Instead of making two separate converters for two voltage domains, one SIDO converter twice as large in area is built in order to power up two separate cores. Proposed converter can provide currents to its two outputs by switching back and forth between two modes by means of TDM. The ratio of the power consumptions of the two cores set the relative time the converter should spend providing current to each of its outputs. For example, supply line of the core with larger power requirement will be charged by the converter more often than that of the core with smaller power requirement. Maximum power obtainable from a converter is proportional to the flying capacitor, which is again proportional to the area of the converter. Hence, in an extreme case, in which one core is completely idle, a SIDO converter with TDDM capability can supply current to the active core for the entire period of time.

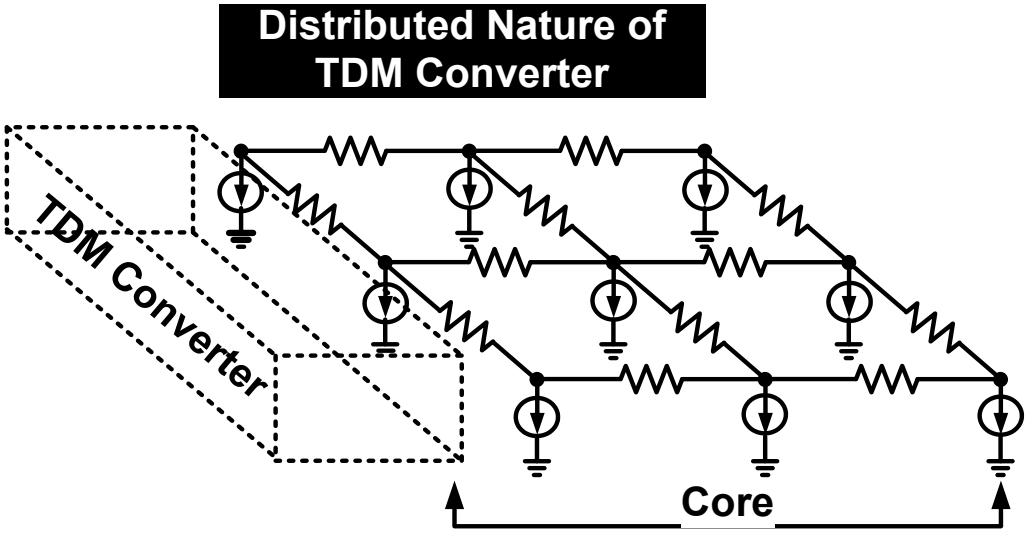


Figure 4.3: Sample converter-core network explaining IR noise reduction using TDM.

Hence, it is able to provide almost twice the output load current to the active core compared to an equivalent conventional single input single output (SISO) converter as shown in Figure 4.2. As output load increases, output voltage of a converter reduces. However, in an unbalanced load scenario, a TDDM

SIDO converter can spend most of its time providing current to the active core. As a result, output voltage of a TDDM converter experiences much less noise compared to an equivalent SISO converter. It has been shown in Figure 4.2 bottom. Alternatively, for a constant output voltage, a TDDM SIDO converter under unbalanced load scenario is capable of dumping more current to the active output as compared to a conventional SISO converter. Hence, maximum output power for one output of the TDDM SIDO converter is larger than which can be obtained from a SISO converter.

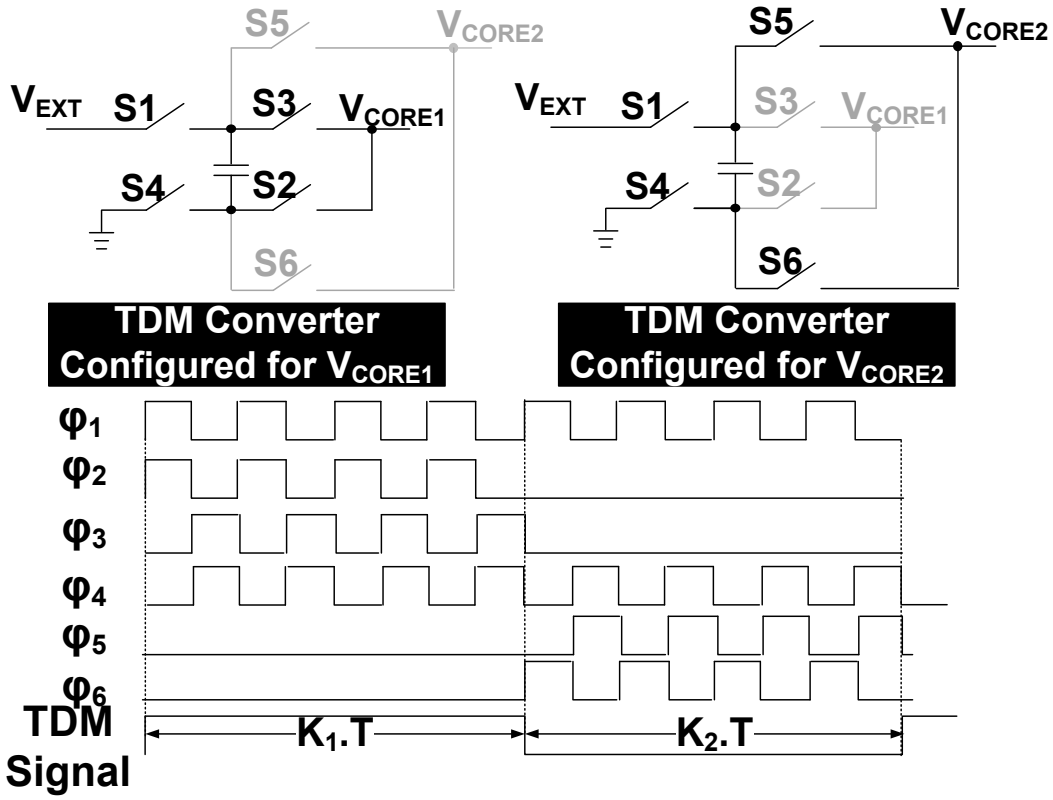


Figure 4.4: Switch schematic of TDDM converter for 1:2 modes (top). Timing diagrams for the switches (bottom).

A secondary advantage of TDDM SIDO converter is that it is distributed in nature. As shown in Figure 4.3, distributed TDDM converter interacts with the microprocessor cores through more number of access

points as compared to a lumped one. As a result of that IR noise seen by the cores tends to be smaller with a TDDM converter than with a conventional SISO converter.

4.2 Analytical Expression of Steady-state Current of a TDDM Converter with 1:2 Voltage Conversion Ratios

A TDDM SIDO converter has to provide power to two cores simultaneously. The amount of time the converter spends on one output determines power delivered by the converter to that particular output. To understand this point, please consider the following case where a single converter generates two outputs with voltage conversion ratio of 1:2, as shown in Figure 4.4. From this figure, we see that the converter down-converts its input voltage V_{EXT} to V_{CORE1} and V_{CORE2}, which will be used as the supply voltage of the cores. Left top figure highlights the switches and the capacitor that are responsible for providing power to the core running at V_{CORE1}. On the other hand, top right figure highlights the configuration corresponding to providing power to V_{CORE2}. In the subsequent analysis, clock signals corresponding to switches S_i are named as φ_i. From the timing diagram of Figure 4.4, we see that during the charging phase when the converter is supplying current to core1, switches S1 and S2 are closed and all other switches remain open. Hence, in this phase, the capacitor is connected between V_{EXT} and V_{CORE1}. During discharging phase of the capacitor for the same output, switches S3 and S4 remain closed, and all other switches remain open. This configuration connects the capacitor between V_{CORE1} and GND. In a similar fashion, during charging phase of the capacitor for output V_{CORE2}, switches S1 and S6 are closed while all other switches remain open. Finally, during the discharging phase of the capacitor for output V_{CORE2}, switches S5 and S4 are kept closed and all other switches remain open. Clock signals that drive these switches are shown in the timing diagram of Figure 4.4. From the timing diagram, we see that the converter spends 4 clock cycles on V_{CORE1}, and 4 clock cycles on V_{CORE2}. Thus, the converter spends same amount of time on each output. From chapter 3, we see that for a single input single output converter, output current is given by the relation

$$i_{AVG\_TOTAL} = 2 \cdot f_{sw} \cdot (V_{EXT} - 2 \cdot V_{CORE}) \cdot \frac{1-e^{-\frac{T}{4RC}}}{1+e^{-\frac{T}{4RC}}}.$$

Hence, for the converter of Figure 4.4, average output current will be $i_{AVG\_SIDO} = \frac{1}{2} \cdot i_{AVG\_TOTAL}$.

In the timing diagram of Figure 4.4, we have shown a waveform for TDDM signal. This TDDM

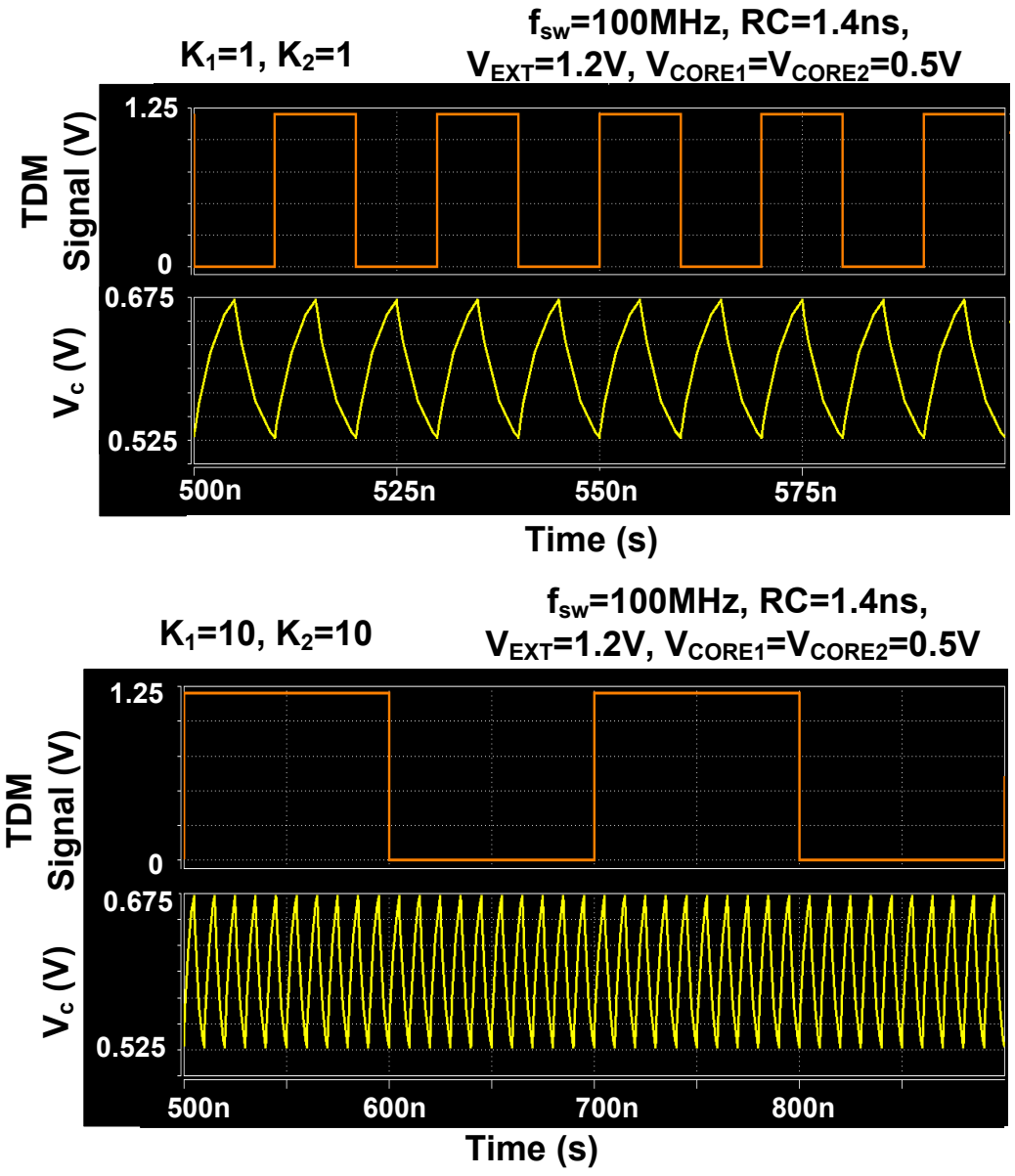


Figure 4.5: Simulation waveforms showing minimal cross-regulation of SIDO converter with equal voltage conversion ratios.

signal denotes how much time the converter spends on each of its inputs. Depending on the respective load condition of its two outputs, the converter can modulate the period of the TDDM pulse signal. If the ratio of loads of core1 and core2 is K_1/K_2, then the converter would spend K_1.T cycles and K_2.T cycles on core1 and core2, respectively. Hence, the average load current supplied to core1 and core2 would be

$i_{AVG\_CORE1} = \frac{K_1}{K_1+K_2}.i_{AVG\_TOTAL} = D.i_{AVG\_TOTAL}$, and $i_{AVG\_CORE2} = \frac{K_2}{K_1+K_2}.i_{AVG\_TOTAL} = (1-D).i_{AVG\_TOTAL}$, respectively, where D is the duty cycle of the TDDM signal shown in Figure 4.4. However, if one core becomes idle, maximum current the converter would be able to provide to its active core would be equal to i_{AVG_TOTAL}.

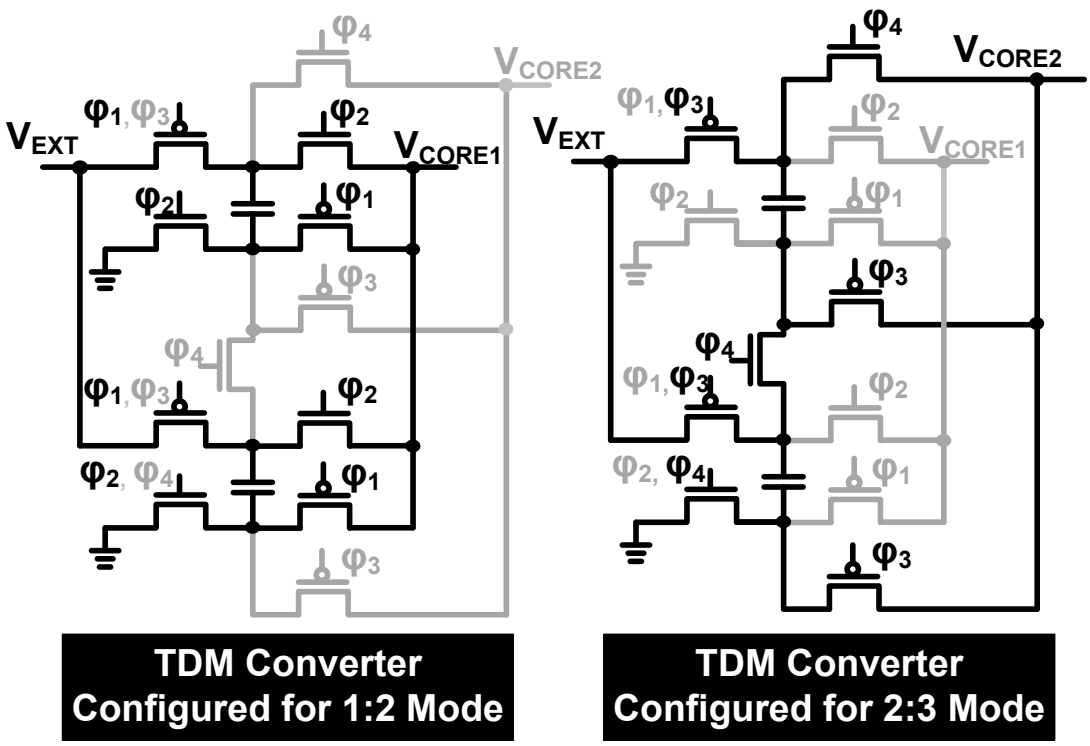


Figure 4.6: Switch schematic of TDDM converter for 1:2 and 2:3 modes.

4.3 Cross-regulation in TDDM Converter

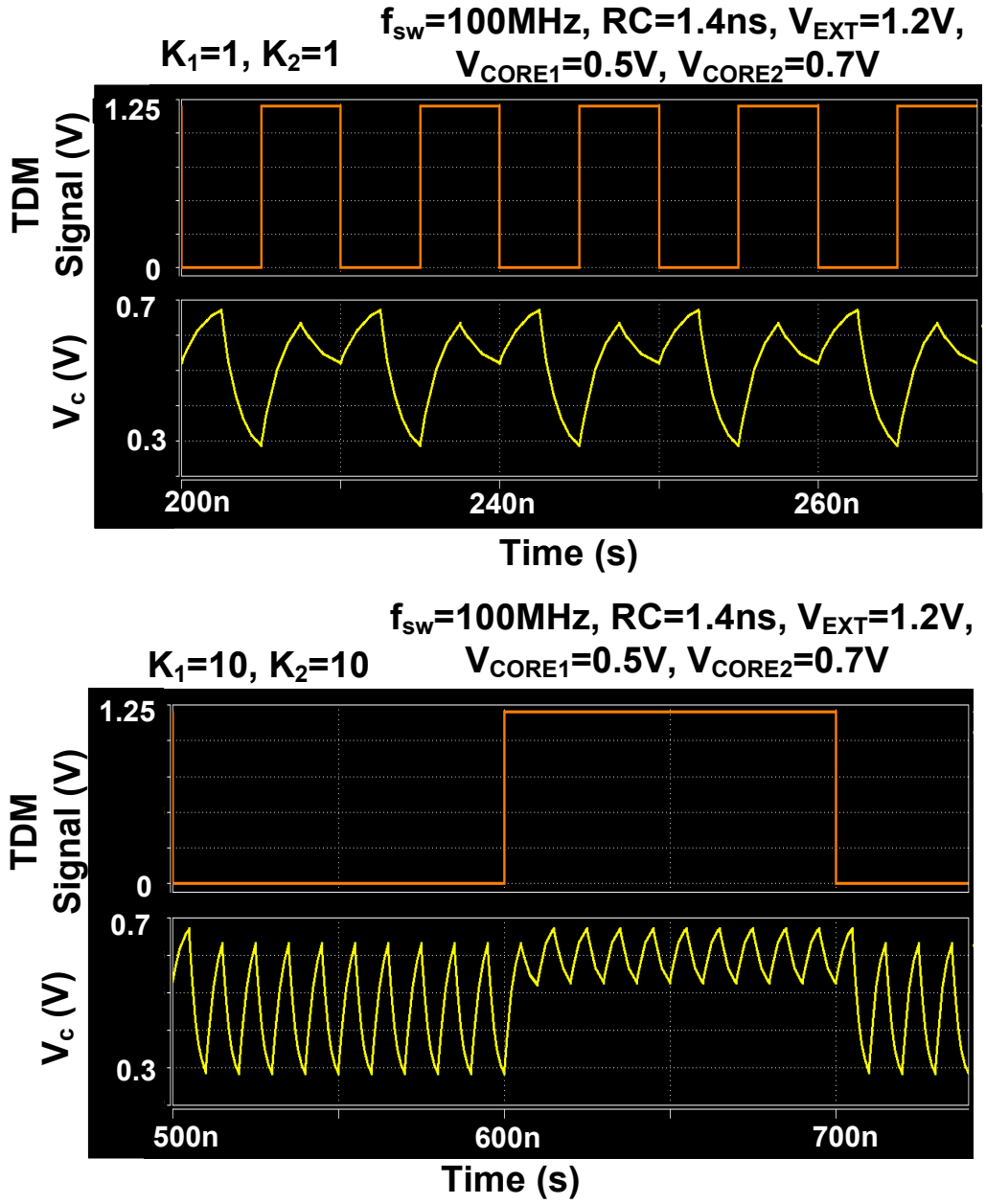


Figure 4.7: Simulation waveforms showing significant cross-regulation of SIDO converter with unequal voltage conversion ratios.

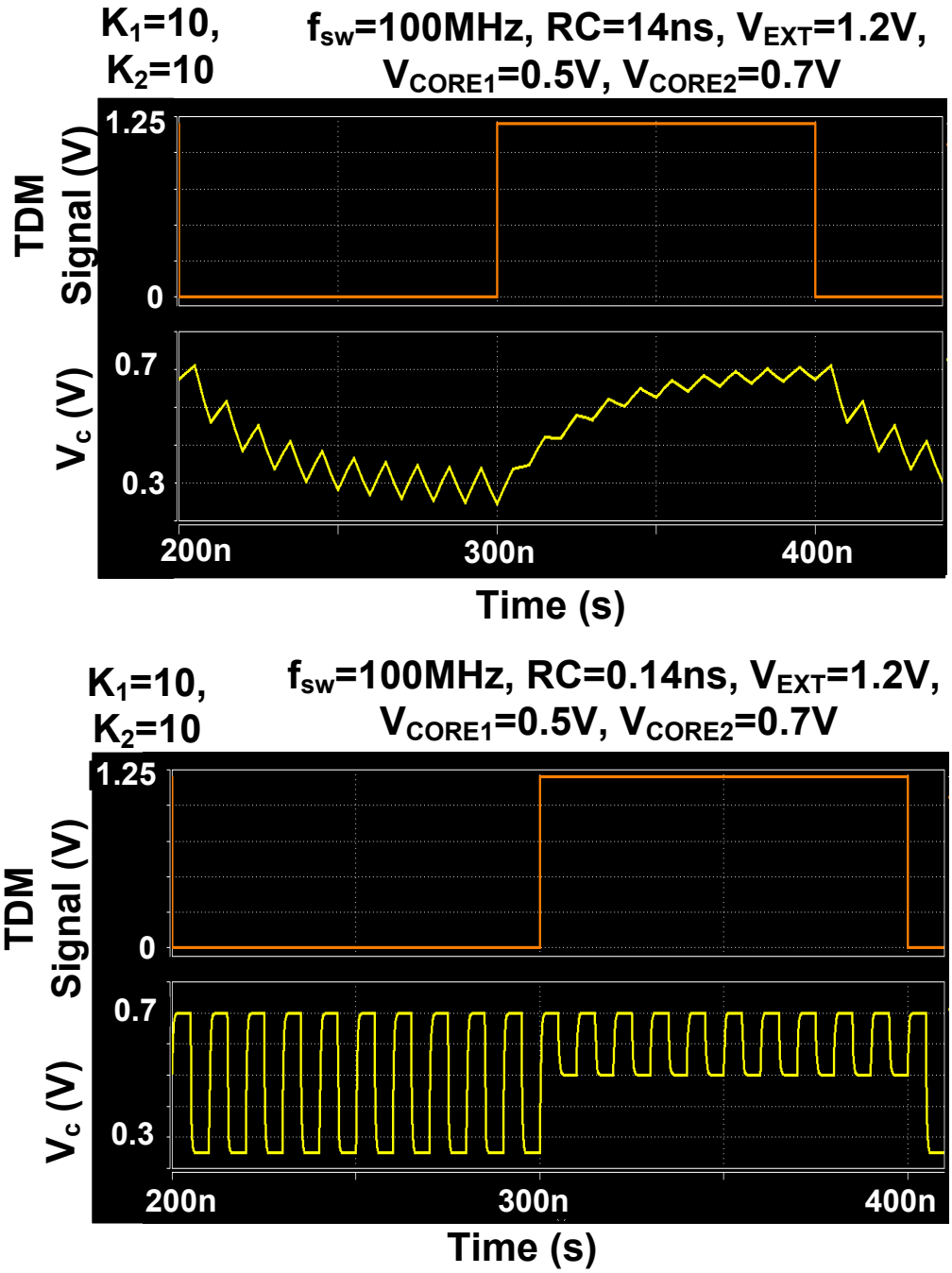


Figure 4.8: Simulation waveforms showing significant cross-regulation for RC~T (top), and negligible cross-regulation for RC<<T (bottom).

TDDM SIDO converter suffers from cross-regulation at steady state. When both voltage conversion ratios of the converter are same as the case in Figure 4.4, steady-state cross-regulation

is minimal as the voltages across capacitor in the steady-state for both the outputs are same. This can be verified from the simulation results shown in Figure 4.5. In this simulation, we assume f_{sw}=100MHz, RC=1.4ns, V_{EXT}=1.2V and V_{CORE1}=V_{CORE2}=0.5V. In this case, the converter generates 1:2 ratios for both outputs. From Figure 4.5 top, we see that K_1=K_2=1, i.e. converter spends one cycle on one output before it provides current to the other output. For both output modes, voltage across capacitor switches between 0.525V and 0.675V. If on the other hand, the converter is allowed to spend 10 consecutive cycles on one output before it can switch to the other output, i.e. K_1=K_2=10, voltage across the capacitor will still change between 0.525V and 0.675V during both output modes. Hence, voltage across capacitor is independent of the number of consecutive cycles the converter spends on its output. Because of this behavior, cross-regulation is negligible during stead-state of the converter when both the voltage conversion ratios are same.

Now, let us take a look at the case where the converter delivers power to its two outputs having two different voltage conversion ratios. For this particular example, we assume voltage conversion ratios of 1:2 and 2:3 as shown in Figure 4.6. We assume the converter spends equal amount of time on both the outputs. However, depending on how many consecutive cycles the converter spends on each output, the amount of steady-state cross-regulation will vary. In other words, the converter needs to spend several clock cycles on one output before the cross-regulation effect from the other output becomes negligible. It has been shown in Figure 4.7. From Figure 4.7 top, we see significant cross-regulation effect on capacitor voltage. It is due to the fact that converter serves its outputs cycle-by-cycle basis. However, when each output is served on ten consecutive cycles, cross-regulation minimizes, as we can see from Figure 4.7 bottom. The number of cycles the converter needs before the cross-regulation in steady-state becomes negligible depends on the RC time constant and f_{sw}. If RC time constant is higher compared to the clock period T, it takes several clock cycles for the capacitor voltage to reach steady-state. That is

why, if the converter spends only one cycle on each output, capacitor voltage cannot reach its steady-state value corresponding to each output, resulting in significant cross-regulation. However, smaller time constant compared to T ensures faster charging-discharging of the capacitor. As a result, cross-regulation minimizes. The effect of time-constant on cross-regulation is shown in Figure 4.8. From Figure 4.8 top, we see that the when RC=14ns, and T=10ns, there is significant cross-regulation between two outputs of the converter. We see that although the converter spends 10 cycles consecutively on both the outputs, the capacitor voltage is unable to reach the steady-state value in either output mode, and severe cross-regulation occurs. When RC is reduced to 0.14ns, and T is kept constant at 10ns, capacitor charging-discharging time reduces significantly, resulting in negligible cross-regulation between the outputs, as can be seen from the capacitor voltage of Figure 4.8 bottom.

Effect of time constant on cross-regulation for a SIDO converter supporting two different voltage conversion ratios can be analytically understood with the help of capacitor charging-discharging equations. We assume a converter with switch configurations as shown in Figure 4.6. This converter supports 1:2 and 2:3 voltage conversion ratios. We assume that all switches have equal ON resistance of R, and each of the flying capacitors of the converter has a capacitance value of C. Let us also assume that the converter starts providing current to the output with voltage conversion ratio of 2:3 with an initial voltage of V_{c0}. Starting with this initial voltage, capacitor charges to $V_{Charge\_1} = (V_{EXT} - V_{CORE2}) + \left(V_{C0} - (V_{EXT} - V_{CORE2})\right).e^{-\frac{T}{4RC}}$ after a time of T/2. In the discharging mode of 2:3 output, capacitor discharges to a value of $V_{Discharge\_1} = \frac{V_{CORE\ 2}}{2} + \left(V_{C0} - \frac{V_{CORE\ 2}}{2}\right).e^{-\frac{2T}{3RC}}$ at the end of the discharging cycle, i.e. after a time of T. A general solution, which expresses the voltage across the capacitor at the end of the charging phase of 2:3 after the converter spends n consecutive cycles on this output, can be given by :

$$V_{Charge\_n} = V_{EXT}.\left(1 - e^{-\frac{3T}{12RC}} + \left(e^{-\frac{11T}{12RC}} - e^{-\frac{14T}{12RC}}\right).\sum_{i=2,n>1}^{n} e^{-\frac{11(i-2)T}{12RC}}\right)$$

$$+ V_{CORE2}.\left(-1 + e^{-\frac{3T}{12RC}}.e^{-\frac{11(n-1)T}{12RC}}\right.$$

$$\left. + \frac{3}{2}\left(e^{-\frac{3T}{12RC}} - e^{-\frac{11T}{12RC}}\right)\sum_{i=2,n>1}^{n} e^{-\frac{11(i-2)T}{12RC}}\right) + V_{C0}.e^{-\frac{3T}{12RC}}.e^{-\frac{11(n-1)T}{12RC}}$$

In a similar fashion, the voltage across capacitor at the end of the discharge cycle after the converter spends n consecutive cycle on 2:3 output can be given by the following equation:

$$V_{Discharge\_n} = V_{EXT}.\left(e^{-\frac{8T}{12RC}} - e^{-\frac{11T}{12RC}}\right).\sum_{i=1}^{n} e^{-\frac{11(i-1)T}{12RC}}$$

$$+ V_{CORE2}.\left(\frac{1}{2} - \frac{3}{2}\left(e^{-\frac{8T}{12RC}} - e^{-\frac{11T}{12RC}}\right)\sum_{i=1}^{n} e^{-\frac{11(i-1)T}{12RC}}\right) + V_{C0}.e^{-\frac{11nT}{12RC}}$$

From the last equation, we find that if the effect of initial voltage across the capacitor, V_{C0}, has to be 1% on $V_{Discharge\_n}$, then $0.01 > e^{-\frac{11nT}{12RC}}$, i.e. $RC < \frac{n}{5.023}T$. Thus, this equation relates RC, T and the

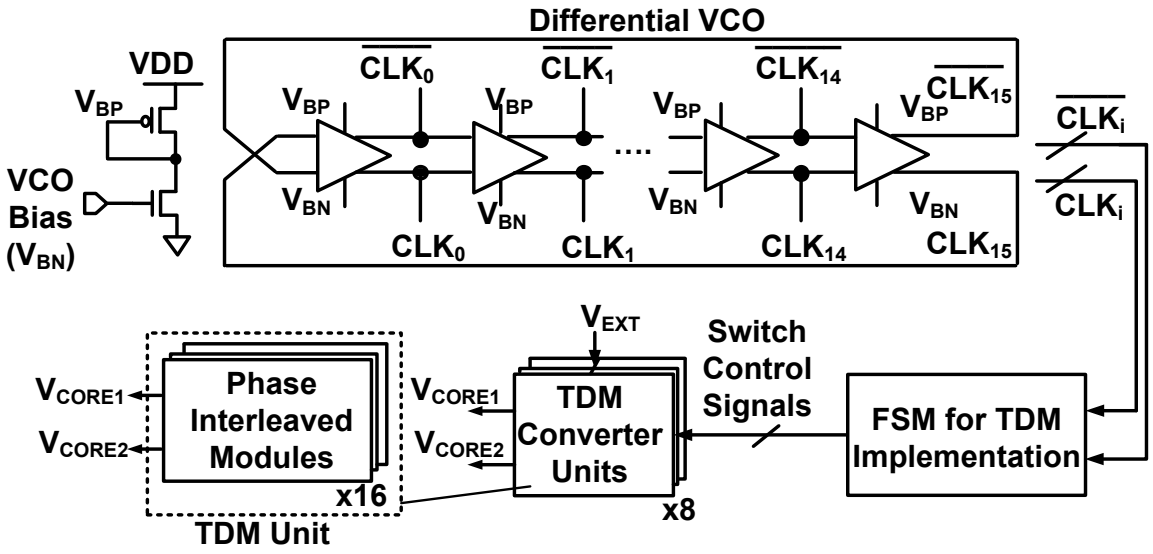


Figure 4.9: Block diagram of the TDDM converter.

number of clock cycles n before the cross-regulation effect becomes negligible.

4.4 Test Chip Implementation of TDDM Converter

Block diagram of the proposed TDDM converter and the cores have been shown in Figure 4.9. The converter, which generates two output voltage levels (V_{CORE1} and V_{CORE2}) with no-load output voltage to input voltage (V_{EXT}) ratios of 1:2 and 2:3, is sandwiched between two cores. The TDDM converter is divided into eight smaller units which are connected in parallel and distributed along the side of the core. For minimizing ripple at the output, each of these eight units is comprised of sixteen multiphase-interleaved stages. A voltage-controlled oscillator (VCO), whose frequency can be controlled by an external bias (V_{BN}), is used to generate 16-phase non-overlapping clock signals. Clock signals are then fed to the finite state machine (FSM), which sends control signals to the converter switch matrix. TDDM parameters such as TDDM signal frequency and duty cycle can be controlled by the FSM unit. Inside each of the cores, there are distributed RC networks and spatially distributed noise injection blocks. In order to generate various current profiles, these noise injection blocks can be turned on and off at precise timing instants using a VCO and clock pattern control units as shown in Figure 4.10.

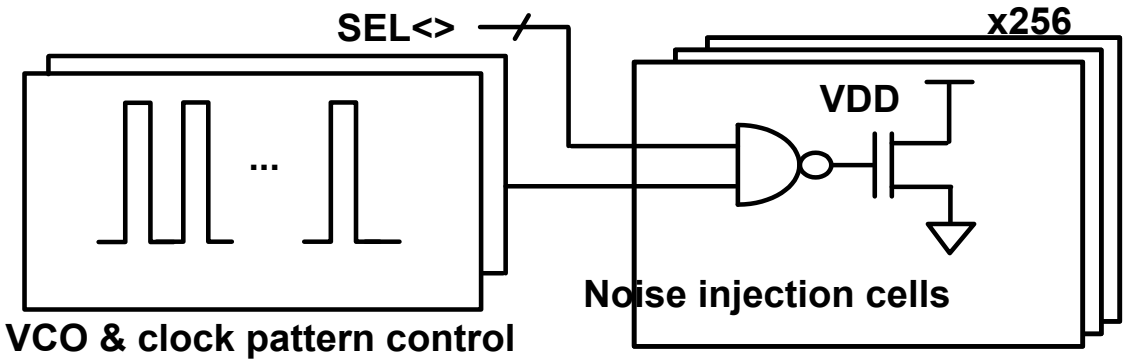


Figure 4.10: Block diagram of a processor core replica.

Schematic of a single phase TDDM converter unit is shown in Figure 4.6. Our proposed converter supports 1:2 and 2:3 modes simultaneously. Switches used in these two configurations are highlighted separately in the two schematics. This single phase TDDM unit uses 2 capacitors and 11 switches in

total. For the same maximum output power, a 1:2 converter would require 2 capacitors and 8 switches, whereas a 2:3 converter would require 2 capacitors and 7 switches. Our TDDM converter can provide power to two outputs simultaneously at the expense of few extra switches. However, because of large minimum distance design rule check (DRC) requirement between two deep trench capacitors in our 32nm process technology, we could place those few extra switches between two capacitors without any

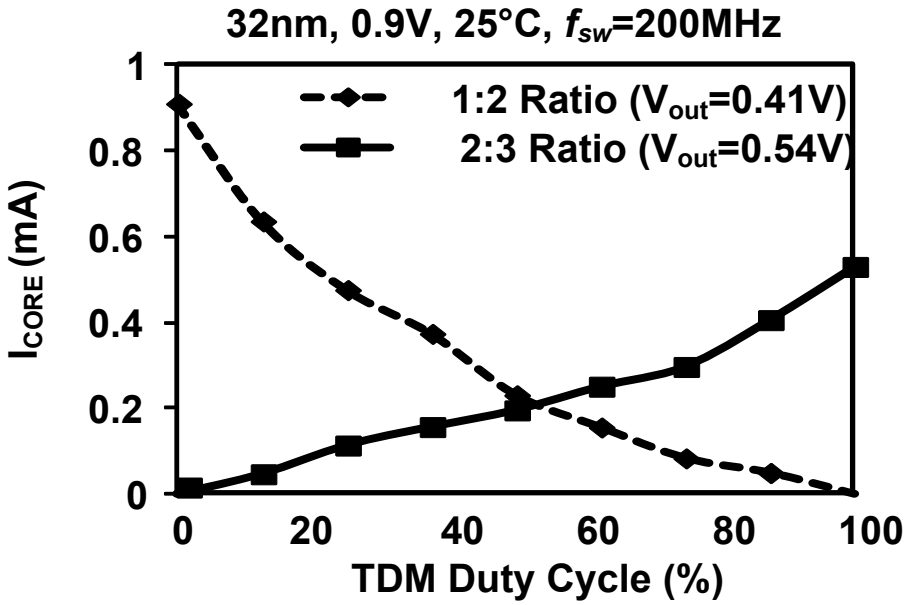


Figure 4.11: Variation in measured load current at a constant load voltage as a function of TDDM duty cycle.

significant increase in the size of the TDDM converter as compared to the size of the 1:2 or 2:3 converters. Since flying capacitors use up more than 70% area of a converter cell, for technologies without deep trench capacitors, we estimate adding few extra switches to build the TDDM converter would increase the area by less than 5-7%. When both the cores are idle, TDDM converter runs at a low frequency to ensure high efficiency at light load, whereas during high activity period of the cores, switching frequency is increased in order to support larger load current. During both these load condition, TDDM pulse is set at 50% duty cycle, which indicates that the converter spends equal time in

both the outputs. During unbalanced load condition, the duty cycle of the TDDM pulse is modulated according to the relative load condition of the two cores.

4.5 Measurement Results and Performance Summary

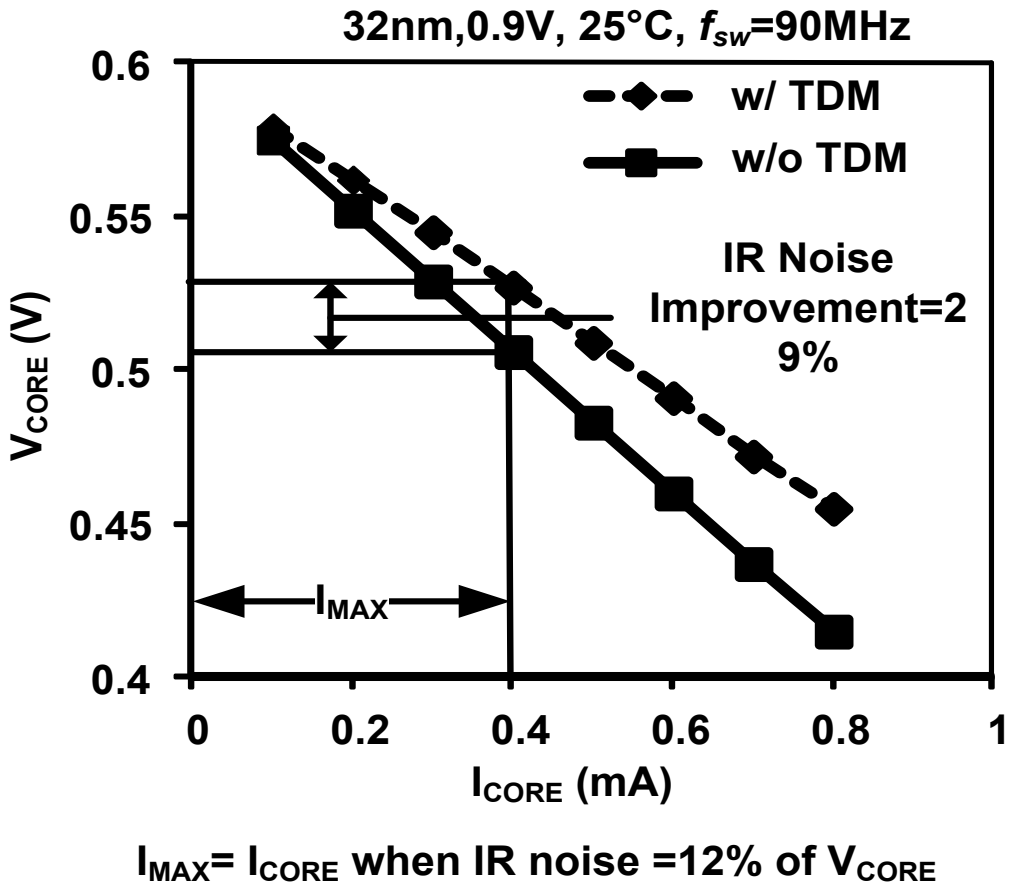


Figure 4.12: IR noise improvement using TDDM converter.

The converter uses deep trench capacitors as flying capacitors, which, to a first order approximation, can be modeled as an ideal capacitance, C_{ideal} and an equivalent series resistance (ESR), R. Because of inherent 3-D nature of deep trench capacitor, very high capacitance can be obtained using a small silicon footprint. Output current of the TDDM converter as a function of duty cycle of the TDDM pulse is shown in Figure 4.11. As duty cycle increases, the converter spends more time to provide current to the

2:3 mode as compared to the 1:2 mode, and hence, output current in 2:3 mode increases and in 1:2 mode decreases. With the condition shown in the figure, TDDM converter can support maximum output currents of 0.54mA and 0.93mA, with 100% and 0% duty cycle, respectively. Because of the distributed nature of the TDDM converter, an improvement of 29% in IR noise has been measured under the

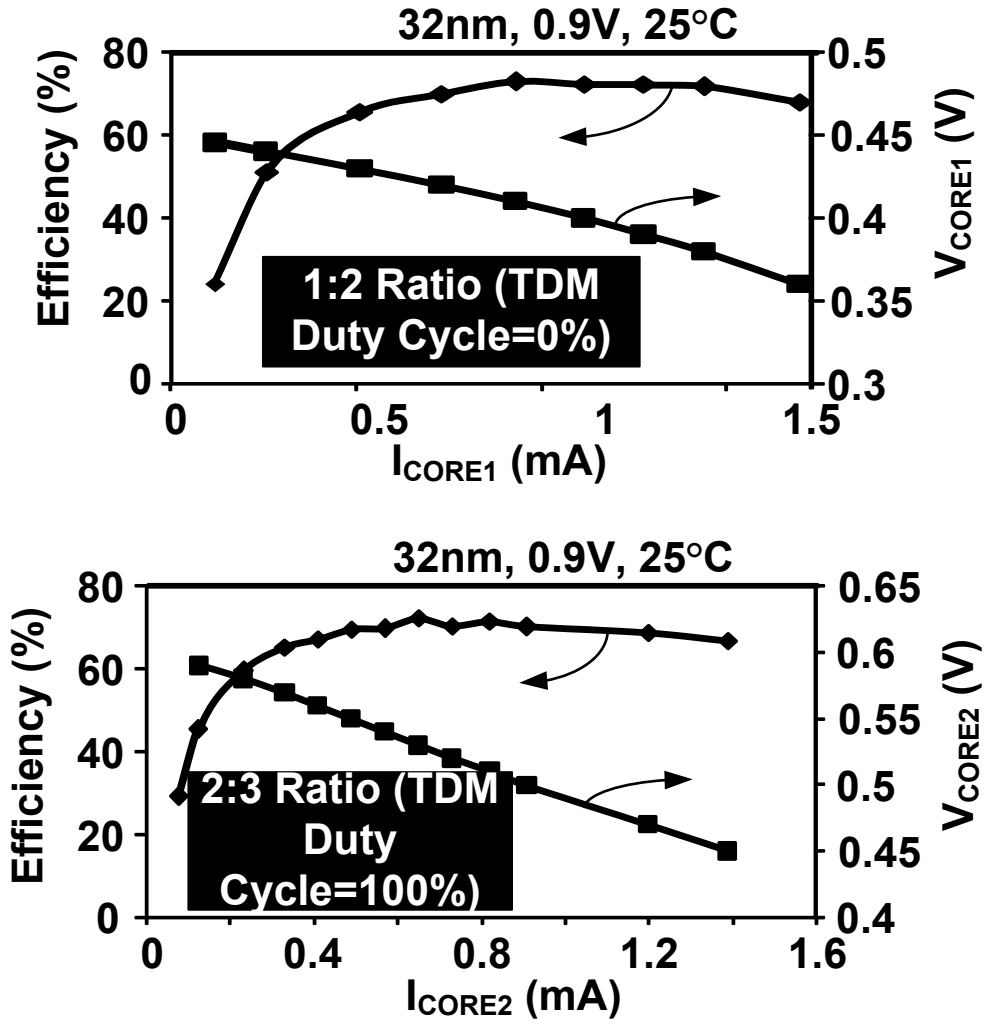


Figure 4.13: Measured efficiency vs. load current and load voltage vs. load current of TDDM converter with 1:2 (top) and 2:3 (bottom) configurations.

conditions shown in Figure 4.12.

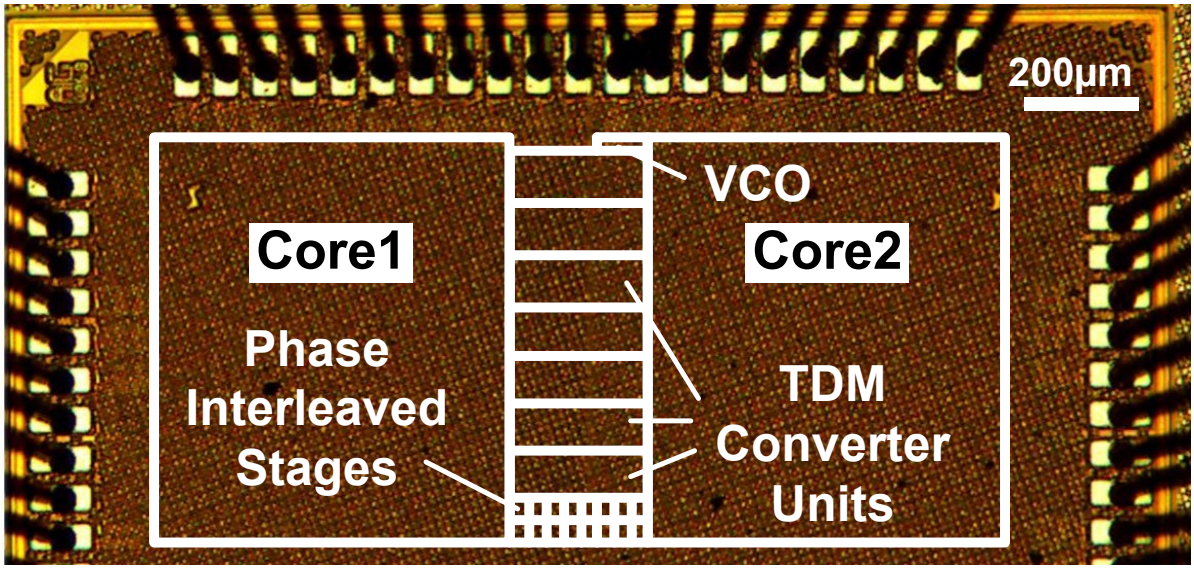


Figure 4.14: Test chip microphotograph.

Figure 4.13 shows efficiency vs. output current along with output voltage vs. output current plots for the TDDM converter for 0% and 100% duty cycle. For V_{EXT}=0.9V, maximum efficiency obtainable from the TDDM converter is 76% using 2:3 ratio (100% TDDM duty cycle), with a corresponding power density of 0.60W/mm^2.

Figure 4.14 shows test-chip microphotograph. Comparison with prior art is shown using table 4.1.

Table 4.1: Performance comparison with current state-of-the-art.

	[44]	[17]	[24]	This work
Technology Node	65nm Bulk	45nm SOI	32nm Bulk	32nm SOI
No. of Voltage Outputs	1			2
Ratios (V_{OUT}/V_{IN})	2:5,1:3 (conv. step-dwn)	2:1,1:2 (conv. step-up/down)	2:1	1:2,2:3 (Distrib. TDM step-dwn)
Flying Capacitor	MOS	Deep Trench	Metal Finger	Deep Trench
Interleaving	18	None	32	16
V_{IN}	3V-4V	2V	1.2V	0.9V
V_{OUT}	1V	0.95V	2V	0.4V-0.6V
IR Noise Improvement	0% (baseline)			29%
Max Efficiency	73%	90%	60%	76% (2:3 TDM)
Power Density at Max Efficiency	0.19W/mm^2	2.19 W/mm^2 @0.95V, 1:2	1.12W/mm^2	0.60W/mm^2 @0.54V, 2:3

5. SYSTEM-LEVEL POWER ANALYSIS WITH ON-CHIP VOLTAGE REGULATORS

5.1 Overview

In order to take full advantage of DVFS, it is important to be able to control supply voltages of the cores of a multi-core processor with a core granularity. Independent control over core voltages makes it possible to reduce the supply voltages of only those cores that are running low-activity jobs without affecting high performance cores, which can keep using higher V_{DD} values as their supply voltage. Hence, with advancement in chip-multiprocessor technology to increase throughput, it became increasingly necessary to be able to control supply voltage of each core in order to reduce total power consumption. Because of smaller component count, faster response time, and the ability to support DVFS on a per-core basis, on-chip integration of power delivery modules has gained momentum over the past few years. There are a few options which are being researched as on-chip power delivery solutions. Among them, on-chip inductor based switching regulators are not suitable for integration because high quality inductors could not be made on-chip [28]. On the other hand, switched-capacitor based on-chip dc-dc converters suffer from relatively low output power density [17]. Even with novel high-density capacitors, switched capacitor dc-dc converters cannot achieve power density required by the micro-processor cores [17]. Although they can be integrated on-chip, this type of converters is not a good candidate for micro-processor applications. As a solution to on-chip power delivery, Intel has introduced a special kind of inductor-based switching converter in its 4th generation Core™ microprocessors (codename Haswell). This converter uses package inductors as the inductors of the switching regulators, and it integrates voltage regulators on die and on package [29]. This kind of switching regulators, which use package inductance in lieu of off-chip inductance, is termed as fully-

integrated voltage regulators (FIVR). IBM has also introduced an on-chip distributed voltage regulator module (VRM) for controlling voltage on a per core basis in its POWER8 processor [30]. In order to estimate power savings using DVFS, it is important to take power loss of the

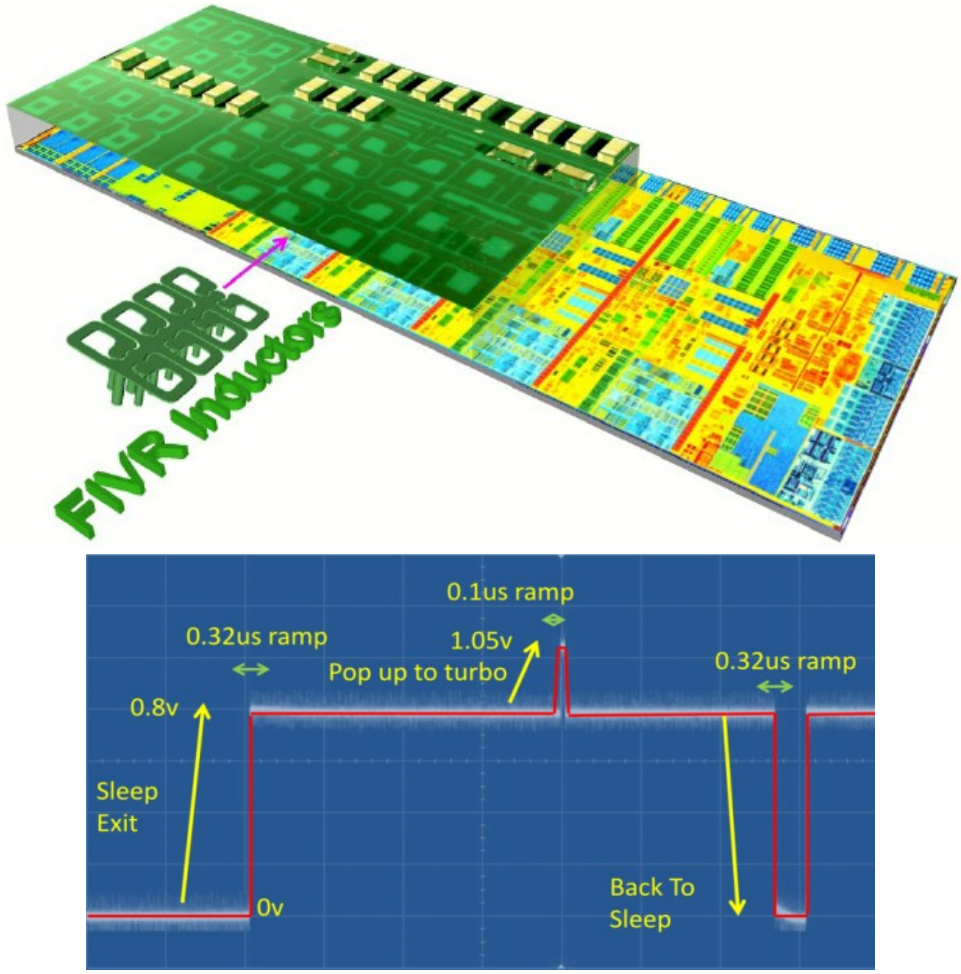


Figure 5.1: Bird's eye view of Intel's Haswell processor die with package inductors for FIVR (above) [29]. Fast DVFS transients enabled by FIVR (below) [34].

voltage regulators into account. Several previous works have attempted to evaluate power/energy benefits of on-chip voltage regulators. For instance, [31] discusses that workload-aware voltage regulator designs can result in system-level energy saving. [32] presents a dynamic reconfiguration of networks that connect voltage regulators to the cores, resulting in system-wide

energy saving. [33] shows that per-core DVFS using on-chip voltage regulation scheme can provide significant system energy reduction. With this knowledge, it becomes necessary to find out which of the on-chip power delivery solutions would cause maximum reduction in system energy/power. Along this line, we concentrate our work on two different state-of-the-art on-chip power delivery solutions (e.g. FIVR and LDO). We consider normalized throughput (introduced in section 5.3) as the performance metric for our 256-core server processor, and use system power consumption as the comparison metric of the two above-mentioned power delivery schemes. The contribution of this project is two-fold. Firstly, it presents a systematic technique to compute average system power of a many-core processor. Secondly, it explores the power-performance design space of the processor using two state-of-the-art power delivery architectures (e.g. FIVR

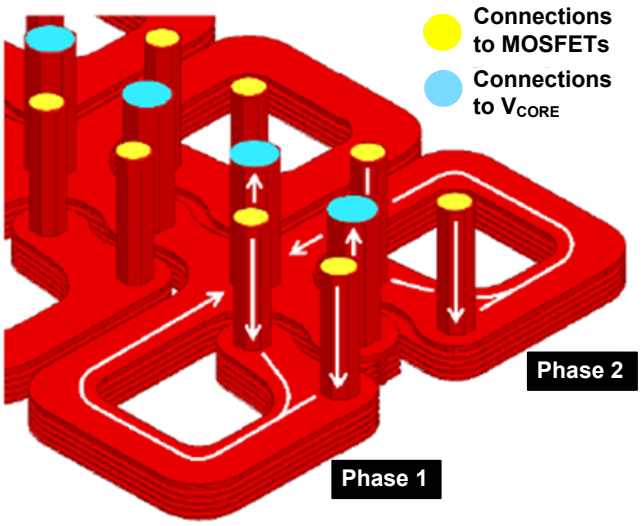


Figure 5.2: 3-D view of two FIVR inductors [29].

and LDO).

5.2 Efficiency Models of Switching Regulators and LDO

Switching voltage regulators are integral components of power delivery systems. Traditional off-chip buck converters typically down-convert off-chip supply voltage to logic voltage used by

microprocessor cores. In this work, in order to support DVFS on a cluster of cores, one more level of voltage conversion has been assumed to take place on-chip, between off-chip buck converter and microprocessor cores. We assume that this on-chip power delivery module can be either FIVR or LDO.

5.2.1 Overview of FIVR

FIVR is a synchronous buck converter built on-chip. It can have up to 16 phases. In order keep filter passives small, FIVR has to be operated at relatively higher frequencies (e.g. 140MHz according to [29]). Cascode NMOS and PMOS power switches built in 22nm Intel's logic process can handle an input voltage of 1.8V and are distributed across the die. They are placed right above the connections of the package inductors in order to minimize routing cost. Because of the close proximity of the regulator and the circuits, extra bumps can be placed on the circuit, and routing can be done using thick metal layer, which effectively increases power density provided by FIVR. Bottom of the package and the die of Intel's Haswell processors along with FIVR inductors have been shown in Figure 5.1 (above). Very fast voltage ramp times in the order of sub-microseconds can be achieved using a FIVR based DVFS system as shown in Figure 5.1 (below), which is another benefit of on-chip power regulators over off-chip ones.

FIVR inductors are non-magnetic. It has air-core. A 3-D view of FIVR inductor with 2-phases has been shown in Figure 5.2. For decoupling purpose, on-chip MIM capacitors and package ceramic capacitors are used. MIM capacitors provide decoupling from output rail and show good transient characteristics. On the other hand, both package ceramic capacitors and on-chip MIM capacitors are used to provide decoupling from the input rail.

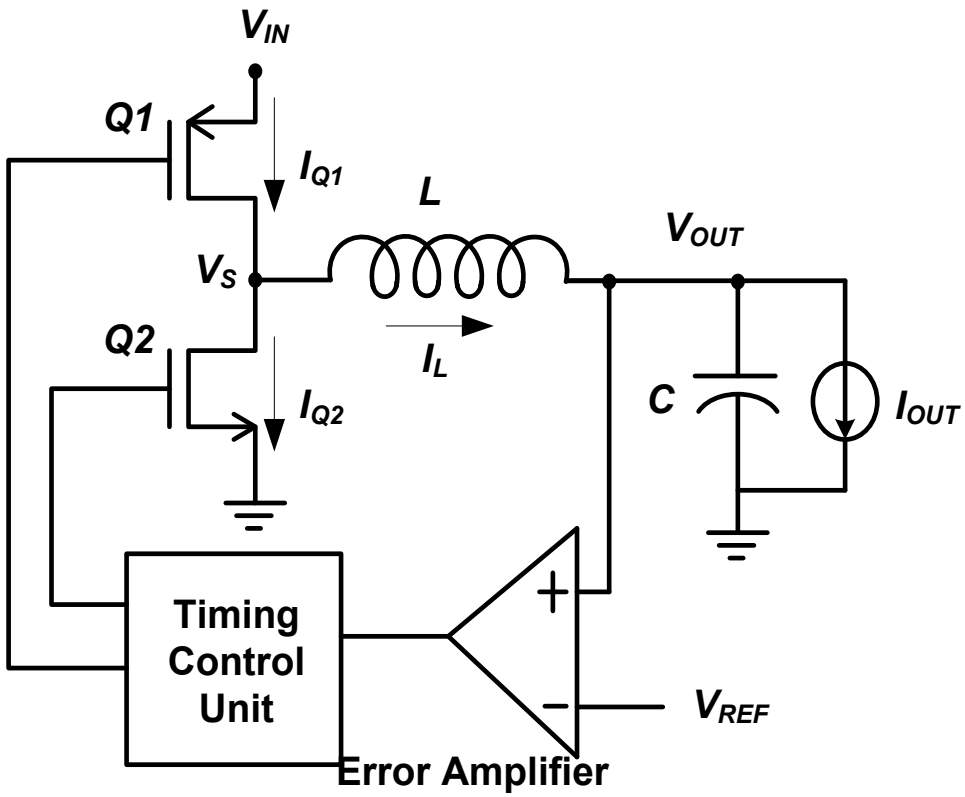


Figure 5.3: Simplified schematic of a step-down switching voltage regulator.

5.2.2 Switching Voltage Regulator Model

Schematic of a generic step-down switching voltage regulator, which can be an off-chip buck converter or an on-chip FIVR, is shown in Figure 5.3. It consists of MOSFET switches (*Q1* and *Q2*), filter network comprising of a filter inductor (*L*) and a capacitor (*C*), and a feedback control loop. Voltage level required by the microprocessor core sets the voltage of the inverting input of a hysteretic comparator. Other input of the comparator is driven by the output of the switching converter. The comparator generates error voltage, which in turn, drives a PFM/PWM controller to generate precise turn-on and turn-off timings of the upper/lower switching MOSFETs, *Q1* and *Q2*. The voltage at the output node of the MOSFETs then drives a low-pass filter formed by *L* and *C*.

Figure 5.4 shows current waveforms of the switching converter through *Q1*, *Q2* and *L*. *Q1* is ON for a time *D.T* during which *Q2* should be OFF, where *T* is the time period of the clock generated by the timing control unit, and *D* is the duty cycle of the clock. *Q2* is kept ON for the remaining

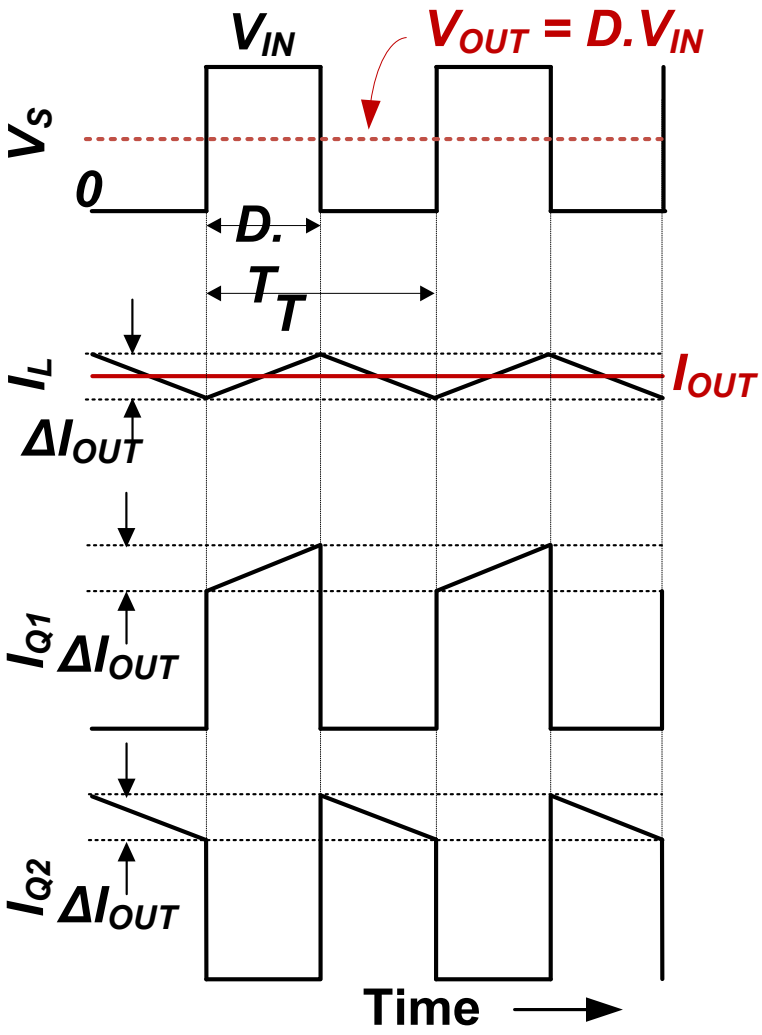


Figure 5.4: Current and voltage waveforms of switching voltage regulator at steady state

of the time period, which is *(1-D).T*. From the current waveforms shown in Figure 5.4, we can see that I_{OUT} is the average output current and ΔI_{OUT} is the inductor current ripple. RMS values of the I_L, I_{Q1} and I_{Q2} can be written as $I_{L,RMS} = \sqrt{I_{OUT}^2 + \frac{\Delta I_{OUT}{}^2}{12}}$, $I_{Q1,RMS} = \sqrt{\frac{V_{OUT}}{V_{IN}} \cdot (I_{OUT}^2 + \frac{\Delta I_{OUT}{}^2}{12})}$ and

$I_{Q2,RMS} = \sqrt{(1 - \frac{V_{OUT}}{V_{IN}}) \cdot (I_{OUT}^2 + \frac{\Delta I_{OUT}{}^2}{12})}$, respectively. Hence, the conduction losses in the switches *Q1* (P_{COND_Q1}) and *Q2* (P_{COND_Q2}), and in the inductor (P_{PAR_L}) can be written as $P_{COND\_Q1} = I_{Q1,RMS}^2 \cdot R_{SW\_Q1}$, $P_{COND\_Q2} = I_{Q2,RMS}^2 \cdot R_{SW\_Q2}$, and $P_{PAR\_L} = I_{L,RMS}^2 \cdot R_{PAR\_L}$, respectively, where R_{SW_Q1} and R_{SW_Q2} are the average on-resistances of switches *Q1* and *Q2*, and R_{PAR_L} is the inductor parasitic resistance. Apart from the conduction loss, the other important loss component is the MOSFET gate drive loss which can be given as $P_{GATE} = C_{GATE} V_{DD}^2 f$, where C_{GATE} is the gate capacitance of *Q1* and *Q2*. Final loss component comes from the power lost in the control circuitry (P_{CTRL}). With these loss components taken into account, power efficiency (η) of a switching regulator can be written as $\eta = \frac{V_{OUT} \cdot I_{OUT}}{V_{OUT} \cdot I_{OUT} + P_{COND\_Q1} + P_{COND\_Q2} + P_{PAR\_L} + P_{GATE} + P_{CTRL}}$.

η attains a peak for a certain load condition. Below this load current, efficiency suffers because of

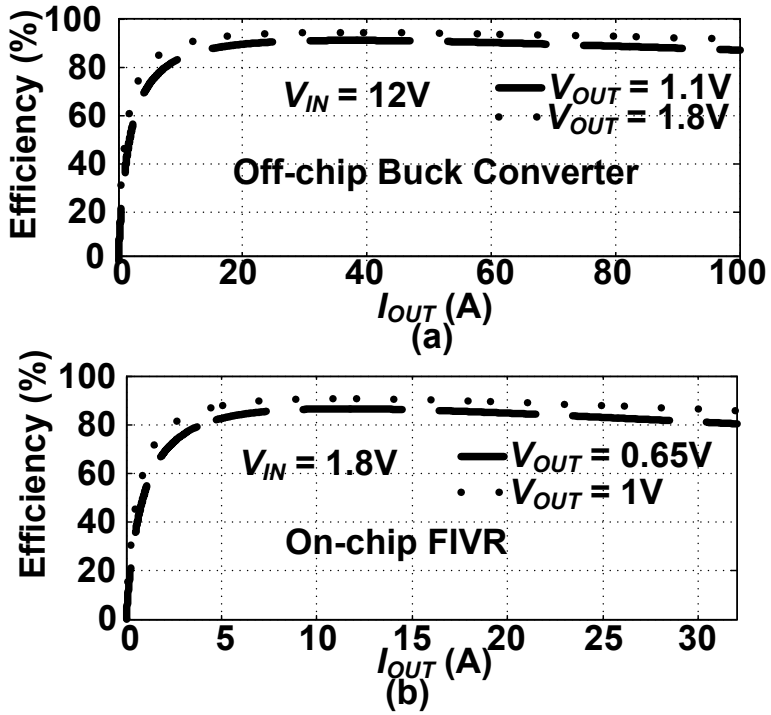


Figure 5.5: Efficiency vs. I_{OUT} for (a) single-phase off-chip buck regulator, and (b) single-phase on-chip FIVR.

load independent gate drive and control circuit loss, and above this load current efficiency drops due to excessive conduction loss. Efficiency vs. load characteristics of off-chip buck converter and on-chip FIVR are shown in Figure 5.5. An off-chip buck converter which sits on a motherboard can typically take 12V from the power supply unit, and down-convert it to the voltage level to be used by the next stage voltage regulator (FIVR or LDO) [35]. FIVR, on the other hand, uses 1.8V as input voltage and generates different voltage levels based on the requirement of the cores, to which the converter is delivering power [29]. As the output voltage of the converter reduces at constant load current, converter efficiency reduces. It can be verified from Figure 5.5.

In order to improve light load efficiency and reduce output voltage ripple, instead of building a

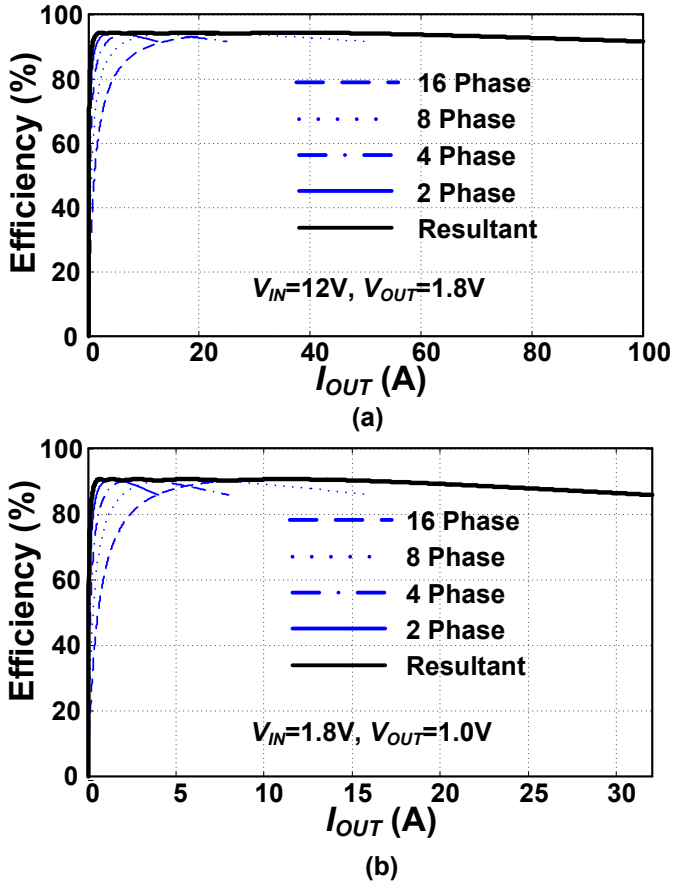


Figure 5.6: Efficiency vs. IOUT for multi-phase (a) buck regulator, and (b) FIVR.

single converter, smaller converter modules are built. Running the converter modules in a phase-interleaved fashion ensures smaller output ripple, and phase dropping at light load ensures improvement in light load efficiency. Typical efficiency vs. I_{out} characteristics of a 16-phase off-chip converter and 16-phase FIVR are shown is Figure 5.6(a), and Figure 5.6(b), respectively.

5.2.3 LDO Model

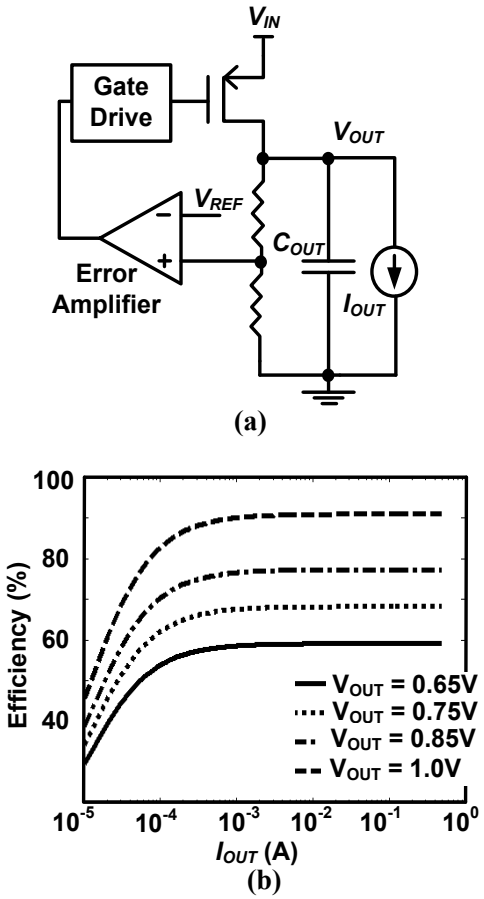


Figure 5.7: (a) Block diagram of an LDO, and (b) Efficiency vs. I_{OUT} of an LDO.

Block diagram of an LDO is shown in Figure 5.6(a). An LDO has an n-type/p-type pass element, which generates regulated output voltage (V_{OUT}) by dropping a portion of the input voltage (V_{IN}) across it. As V_{IN} reduces, or load (I_{OUT}) increases, V_{OUT} starts to drop and is sensed by the error amplifier. Error amplifier then generates larger gate drive to maintain output regulation. In order

that the output is regulated at a proper value, a minimum voltage, known as drop-out voltage of the regulator, has to be maintained across the pass gate. Efficiency of an LDO can be given as

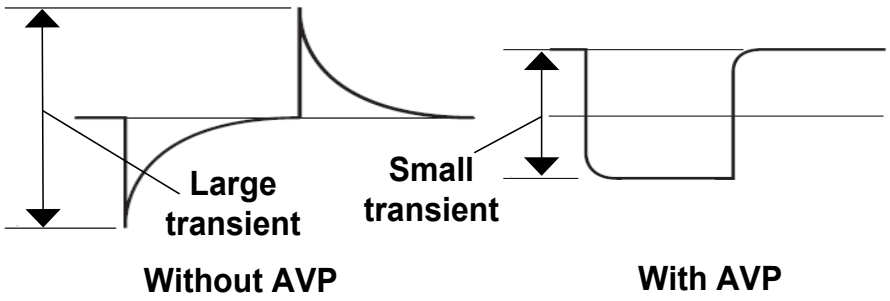


Figure 5.8: Transient response improves with AVP [37].

$$\eta = \frac{V_{OUT} \cdot I_{OUT}}{V_{IN} \cdot (I_{OUT} + I_q)}$$

, in which I_q is quiescent current of the LDO circuitry [36]. As output voltage deviates further from the input voltage, loss in the pass element increases, and efficiency of the regulator reduces. Efficiency of an LDO is also limited by I_q. At light load condition, I_q dominates over the load current, and hence, LDO efficiency drops at light load. Efficiency vs. I_{OUT} characteristics of modeled LDO for a range of output voltages are shown in Figure 5.6(b).

5.2.4 Active Voltage Positioning (AVP)

In order to reduce output ripple during transient, regulation at the output of the converter is not made perfect [37]. At minimum load, V_{OUT} is set at a slightly higher voltage than its nominal value. Regulation is done in such a way that at full-load condition V_{OUT} attains its nominal value. This technique is known as active voltage positioning and is commonly used in voltage regulators in order to reduce transient microprocessor power at the expense of reduced output regulation. Simple waveforms in Figure 5.8 show that AVP reduces the peak-to-peak output excursion. Although this work concentrates on the steady state power consumption of the system, and power consumed during microprocessor transient is out of scope of this work, we still incorporate AVP

into our steady state system power analysis, as AVP modulates steady-state load characteristics of the voltage regulators.

5.3 Power Delivery System Architecture

Figure 5.9 shows the block diagram of the power delivery system used in this work. There is one 16-phase buck regulator sitting outside of the chip on the motherboard. It uses 12V supply and

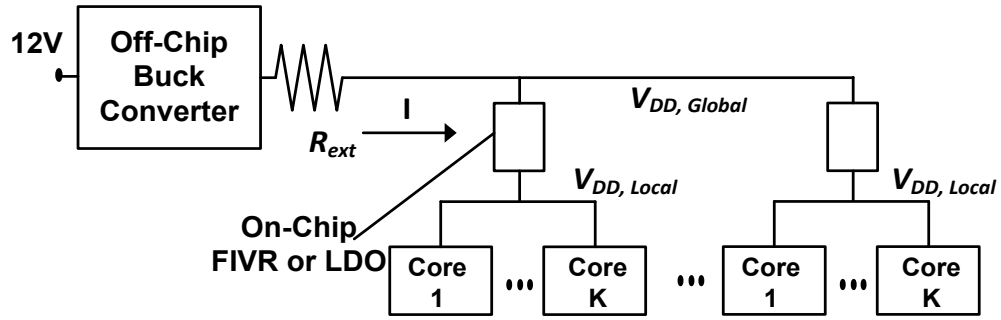


Figure 5.9: Block diagram of power delivery architecture under consideration.

generates output voltage levels to be used by subsequent converter stages. Efficiency vs. load characteristics of this buck converter are shown in Figure 5.6(a). Inside the chip, there are on-chip regulators (FIVR or LDO) and processor cores. IR noise due to external wire and package

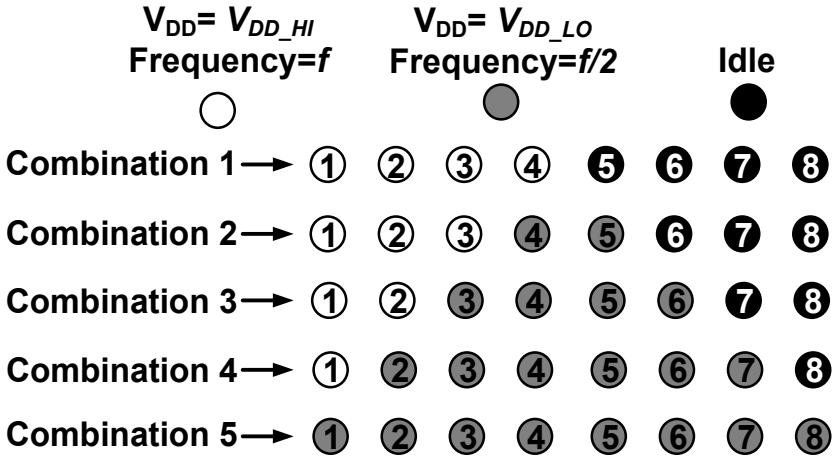


Figure 5.10: Iso-throughput of 0.5 with different power consumption in a multi-core scenario.

resistances is accounted for with a single lumped resistor, R_{ext}, in our simplified model of the power delivery network. We assume that due to R_{ext} , the worst case power loss is approximately 5%, which is typical of current state-of-the-art power delivery networks.

In our analysis, we have assumed a processor with 256 cores. This assumption is in line with the number of cores in several recently developed processors including NVIDIA's GPU accelerator Tesla K80 that has 4992 CUDA cores [38], and Intel's Xeon Phi™ processor that can have up to 61 cores [39]. However, our system power estimation methodology will be applicable to processors with any number of cores. The cores in our hypothetical processor are assumed to be homogeneous in nature, and they can be power gated individually. Supply line of these cores ($V_{DD,\ Local}$) is driven by LDO in case of LDO-based power delivery architecture or by FIVR in case of FIVR-based power delivery architecture. Each of these cores has DVFS capability with maximum and minimum operating frequencies of *f* and *f/2* for corresponding logic V_{DD_min} values of 1V (V_{DD_HI}) and 0.65V (V_{DD_LO}), respectively. These cores are also equipped with a power-down mode for idle state. Although a continuous DVFS scheme would be more useful in terms of

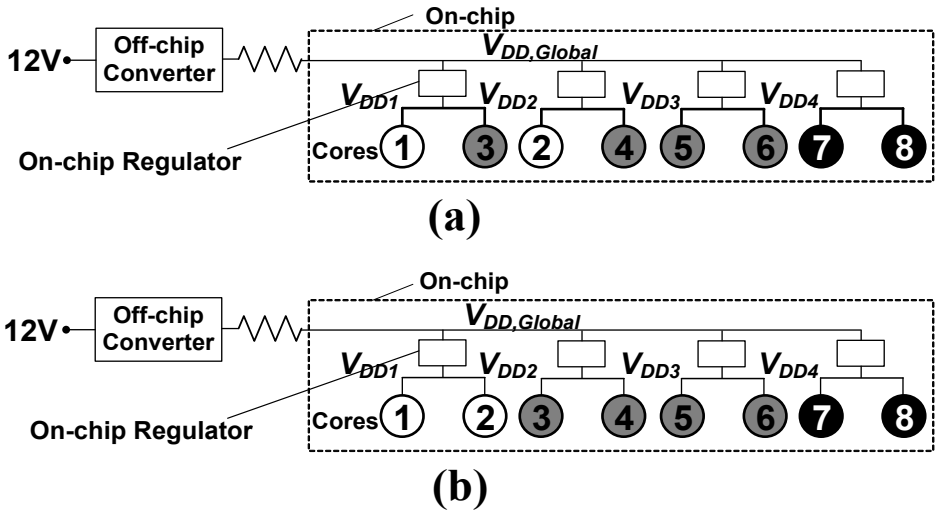


Figure 5.11: Core assignment across power domains: (a) Inhomogeneous VDD, (b) homogeneous VDD.

power savings, its implementation in a 256-core processor might be limited because of synchronization overhead across cores. Also, it can be seen from Figure 5.1 (below) that the supply voltage usually switches between the maximum and minimum levels. Our assumption of two-point DVFS (i.e. 1V or 0.65V) gives us sufficient insight into the total system power consumption with a simple analysis. Please note that 6T SRAM based caches are usually operated under a separate nominal voltage due to read and write margin constraints, and they will require separate voltage regulators. Our analysis is focused on power delivery to the core logic only.

In FIVR based power delivery architecture, we perform the analysis assuming various number of FIVRs (i.e. 16, 128 and 256) present on-chip. For a 256 core processor, it translates to 16, 2 and 1 core per power domain, respectively. We assume future FIVR technology will be able to support more FIVRs per package, and it will be possible to supply a 256-core processor with a single core granularity using FIVRs. Output voltage of FIVR is determined by the activity of the cores powered up by that FIVR. Unless all the cores inside a FIVR domain run at a frequency of *f/2*, voltage of that domain has to be maintained at V_{DD_HI} in order to maintain throughput. However, if all the cores in a power domain can run at a frequency of *f/2*, then that power domain voltage can be set to V_{DD_LO}. Per-core DVFS is possible when number of FIVR increases to 256. Here, in our analysis, we assume no-load and full-load input voltages of FIVR are 1.8V and 1.7V [29]. Efficiency vs. load characteristics of FIVR are shown in Figure 5.6(b).

LDO-based power delivery architecture, on the other hand, can use 256 LDOs to supply power to 256 cores. This is due to the fact that LDOs are inexpensive to build and usually consumes silicon area of the same order as power gates. Because of the presence of 256 LDOs on-chip, per-core DVFS is possible. However, it does not guarantee better total system power than FIVR based architectures because of lower LDO efficiency. In our analysis, we assume an LDO whose efficiency vs. load characteristics are shown in Figure 5.6(b).

5.4 System Power Analysis Methodology

In this section, we describe the methodology to obtain total system power for the power delivery architectures mentioned in the previous section. For our processor, we assume a throughput-oriented architecture like GPU, in which the processor has a lot of inherent parallelism. Because of our choice of throughput as system performance metric, we used power consumption instead of energy consumption as the comparison metric of FIVR and LDO based power delivery architectures. In case all the cores run at maximum frequency, we assume a normalized throughput of 1. However, same throughput can result in different powers consumed by the cores. Figure 5.10 shows various core configurations for normalized throughput of 0.5 in an 8-core configuration. From this figure, we find that, in order to obtain normalized throughput of 0.5, 4 cores can run at frequency *f*, whereas other 4-cores can remain idle. However, this particular combination results in maximum power consumption, and it is equal to $P_{8Core} = 4 \cdot C_{EFF} V_{DD\_HI}^2 f + 4 \cdot P_{Leak} + 4 \cdot P_{Static}$. Here, C_{EFF} is the effective dynamic capacitance of each core including activity factor, P_{Leak} is the leakage power of an active core, and P_{Static} is the static power of an idle core. Lowest possible core power consumption corresponds to the case when all the cores run at a frequency of *f/2*, and it is equal to $P_{8Core} = 8 \cdot C_{EFF} V_{DD\_LO}^2 (\frac{f}{2}) + 8 \cdot P_{Leak}$. Since at normal operating condition, P_{Leak} and P_{Static} are smaller than dynamic power, we find that the latter core combination consumes smaller core power under iso-throughput condition, albeit at the expense of a longer execution time. However, we assume, in a power budget-constrained iso-throughput scenario, the processor might have to sacrifice latency in lieu of smaller power.

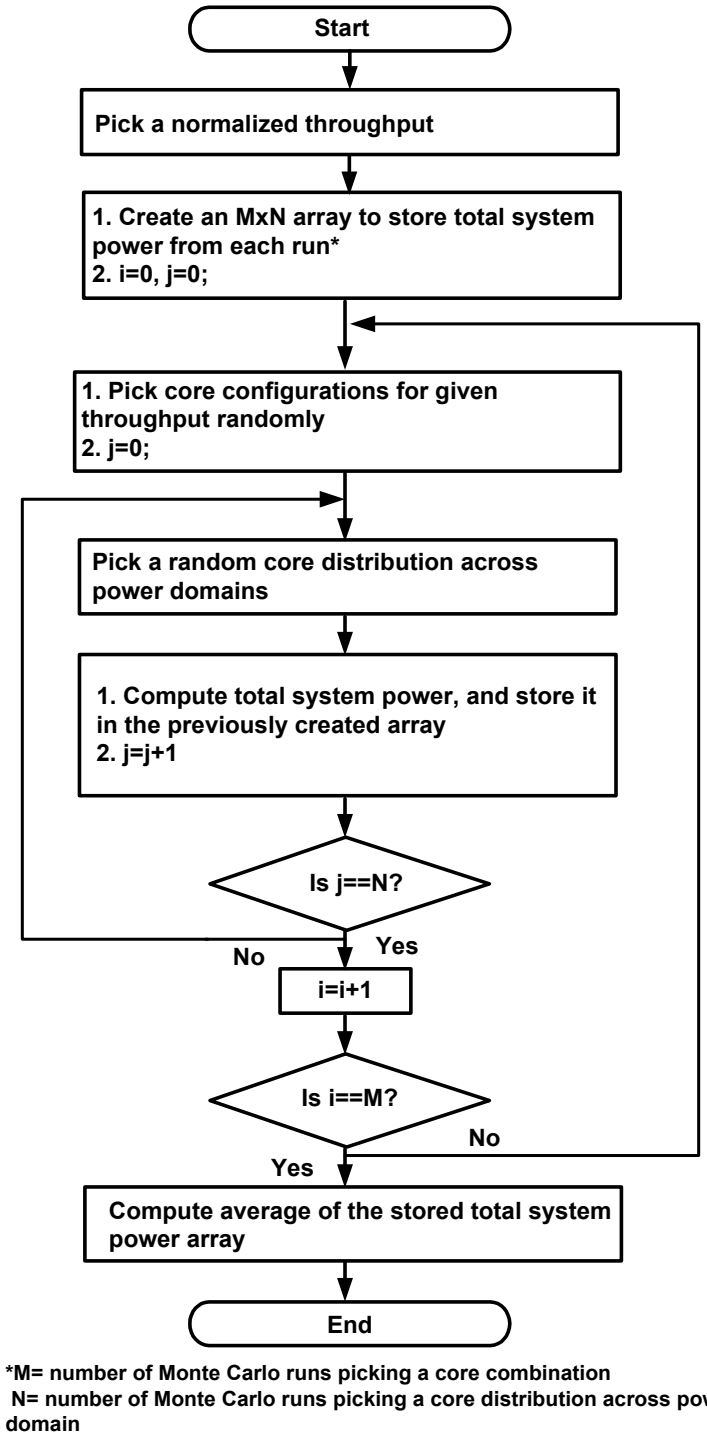


Figure 5.12: Flowchart showing average power computation steps.

In case per-core DVFS is not a viable option, total power consumed by the cores may vary greatly

depending on how the cores are distributed among different power domains. In order to explain this point, we pick combination 3 from Figure 5.10, and distribute the cores across 4 power domains in two different ways as shown in Figure 5.11(a) and 5.11(b).

Figure 5.11(a) shows that all the cores with equal supply voltage requirement have been grouped together in the same power domain. However, this is not the case in Figure 5.11(b). Because of inhomogeneous grouping of the cores in Figure 5.11(b), total power consumed by all cores will be larger compared to the case shown in Figure 5.11(a).

From design space exploration point of view, we use a two-step Monte-Carlo simulation in order

From design space exploration point of view, we use a two-step Monte-Carlo simulation in order

to compute average system power (Flowchart shown in Figure 5.12), and we call it a random scheduling technique. In order to understand how the average system power computation is done with this technique, let us go back to our 8-core processor and pick a normalized throughput of 0.5. For this throughput, we assume the scheduler picks any combination from Figure 5.10

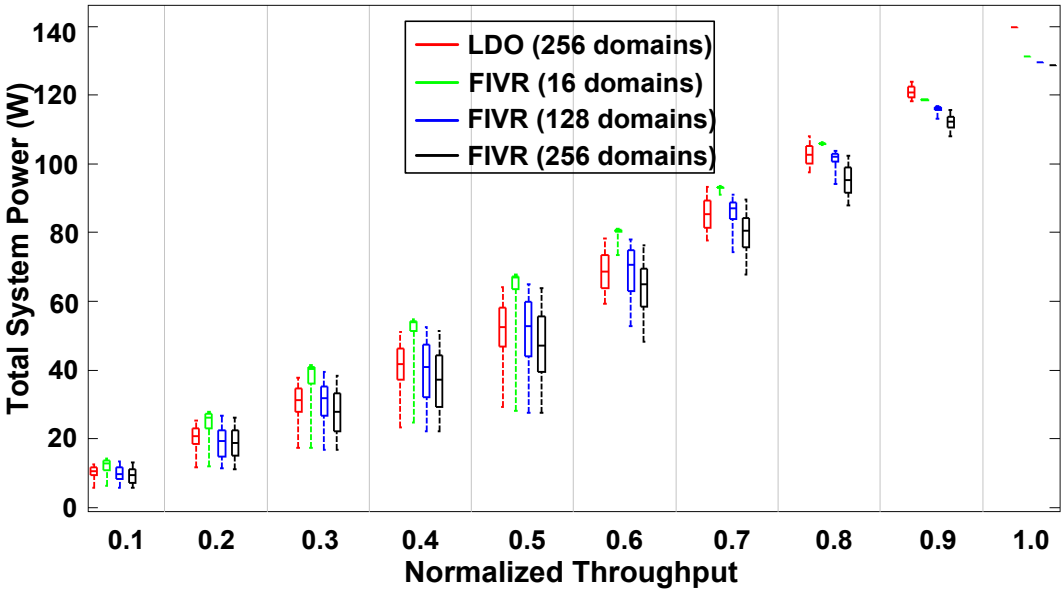


Figure 5.13: Boxplot of power consumption vs. normalized throughput for various power delivery architectures (with process variation).

randomly. We further assume that the cores corresponding to that combination can be distributed across different power domains in a random fashion. Now, if we run Monte-Carlo simulation for this two-step randomization process, and compute an average of the system powers obtained from all the occurrences, we will obtain average system power.

At the same time, we can find out minimum system power using a minimum power scheduler.

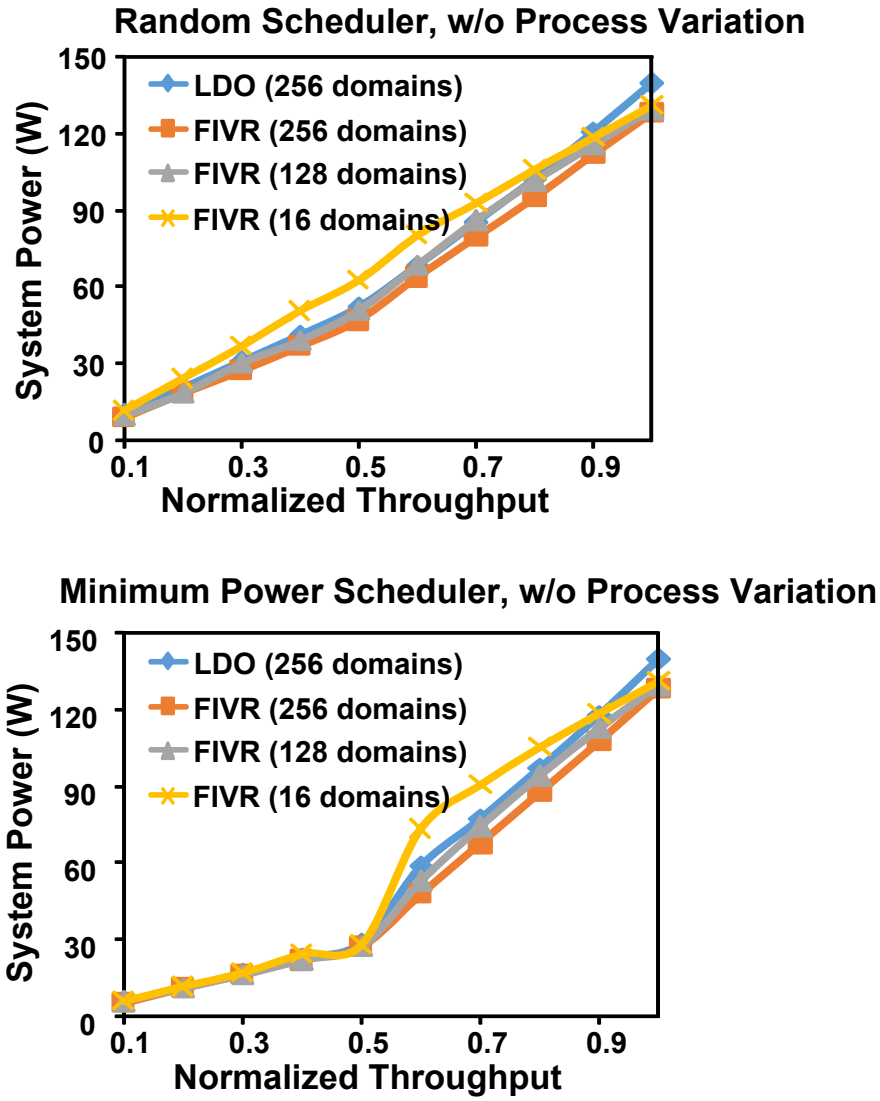


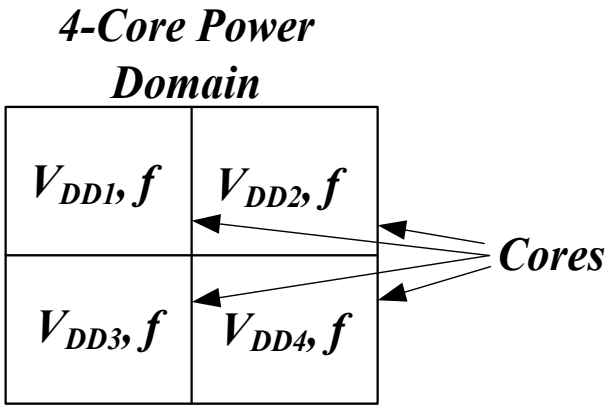
Figure 5.14: System power vs. throughput for various power delivery architectures (w/o process variation).

Our minimum power scheduler will always pick combination 5 from Figure 5.10. Minimum power scheduler will also maximize the number of homogeneous power domains in order to minimize total power consumed by the cores [40].

Once configurations of the cores across power domains are decided, total system power can be obtained by adding the power consumed by cores, on-chip voltage regulators, power distribution network, and off-chip voltage regulator.

5.5 Simulation Results

In our analysis, we assume a hypothetical 256-core processor with (a) 256 power domains for LDO based power delivery architecture, and (b) 16,128 and 256 power domains for FIVR based power delivery architecture. Monte-Carlo simulations were performed at a constant normalized throughput for all the power delivery architectures under consideration. Figure 15 shows a boxplot of system power versus normalized throughput assuming no process variation. This figure indicates that the range of power consumption is maximum for a normalized throughput of 0.5, and it tapers down gradually on either sides of mid-throughput region. It is due to the fact that towards mid-throughput region, number of combinations (as shown in Figure 5.10) to obtain



VDD of the power domain= max(V_{DD1}, V_{DD2}, V_{DD3}, V_{DD4})

Figure 5.15: Slowest core determines supply voltage of the power domain.

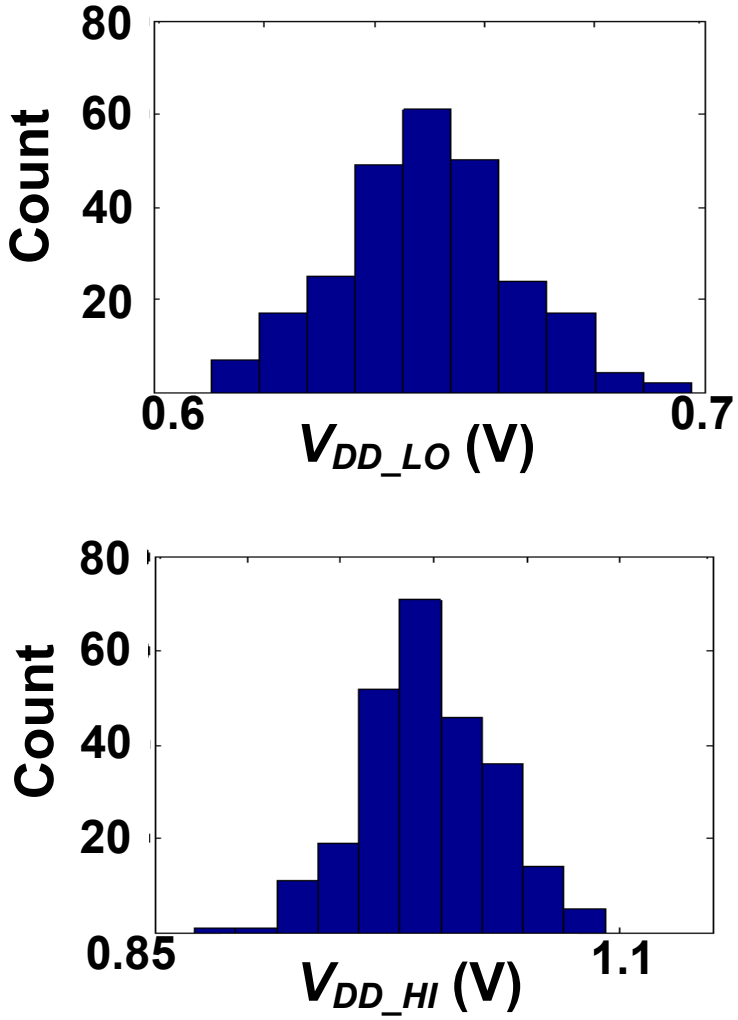


Figure 5.16: V_{DD_min} distribution for 256 cores.

same throughput increases, thereby increasing the power consumption range. For a random scheduler, in order to get average system power for a throughput, we compute an average of the power consumption values for that particular throughput from Figure 5.13. As for a minimum power scheduler, power consumption corresponds to the minimum power point corresponding to each throughput values of Figure 5.13.

Figure 5.14 shows system power plotted against normalized throughput obtained from random scheduler (left) and minimum power scheduler (right). From this figure, we see that the average

system power with LDO is smaller than that with 16 FIVRs because of per-core DVFS capability with LDO (24% less power consumption at normalized throughput of 0.5). However, if the number of power domains using FIVR increases to 128, FIVR architecture becomes comparable to that of LDO because of better FIVR efficiency. Eventually, power consumption with 256 FIVR domains outperforms LDO based architecture by 12% at a throughput of 0.5 for random scheduling technique. Note that perfect homogeneous grouping of cores is always possible when normalized throughput $\leq$ 0.5. Hence, from Figure 5.14, we see that the minimum system power is independent of DVFS architecture for normalized throughput $\leq$ 0.5.

In the presence of process variation, voltage-frequency relations of different cores are different, and power domain supply voltage is determined by the slowest core. It is shown in Figure 5.15. As a result, total dynamic power consumption increases. To a first order approximation, variation in supply voltage is due to variation in transistor threshold voltage [41]. For the sake of analysis, we assume V_{DD_LO} is normally distributed with a mean value of 0.65V and standard deviation of 16mV, and V_{DD_HI} is normally distributed with a mean value of 1.0V and standard deviation of 35mV (Figure 5.16). Figure 5.17 shows total system power vs. normalized throughput using a random scheduler and minimum power scheduler, taking process variation into account. Process variation makes LDO based DVFS approach less attractive than FIVR based one. V_{IN} for the LDOs, which is also $V_{DD,Global}$ in Figure 5.9, has to be determined by the slowest core under process variation. Hence, efficiencies of the LDOs drop. At throughput of 0.5, for random scheduling technique, FIVR with 128 domains burns 10% less power than per-core LDO, and FIVR with 256 domains consume 20% less power than per-core LDO. Finally, in Figure 5.18, we plot average system power vs. number of cores in a processor for a throughput of 0.5. The plot shows that as the number of cores increases, percentage reduction in average system power using per-core FIVR increases.

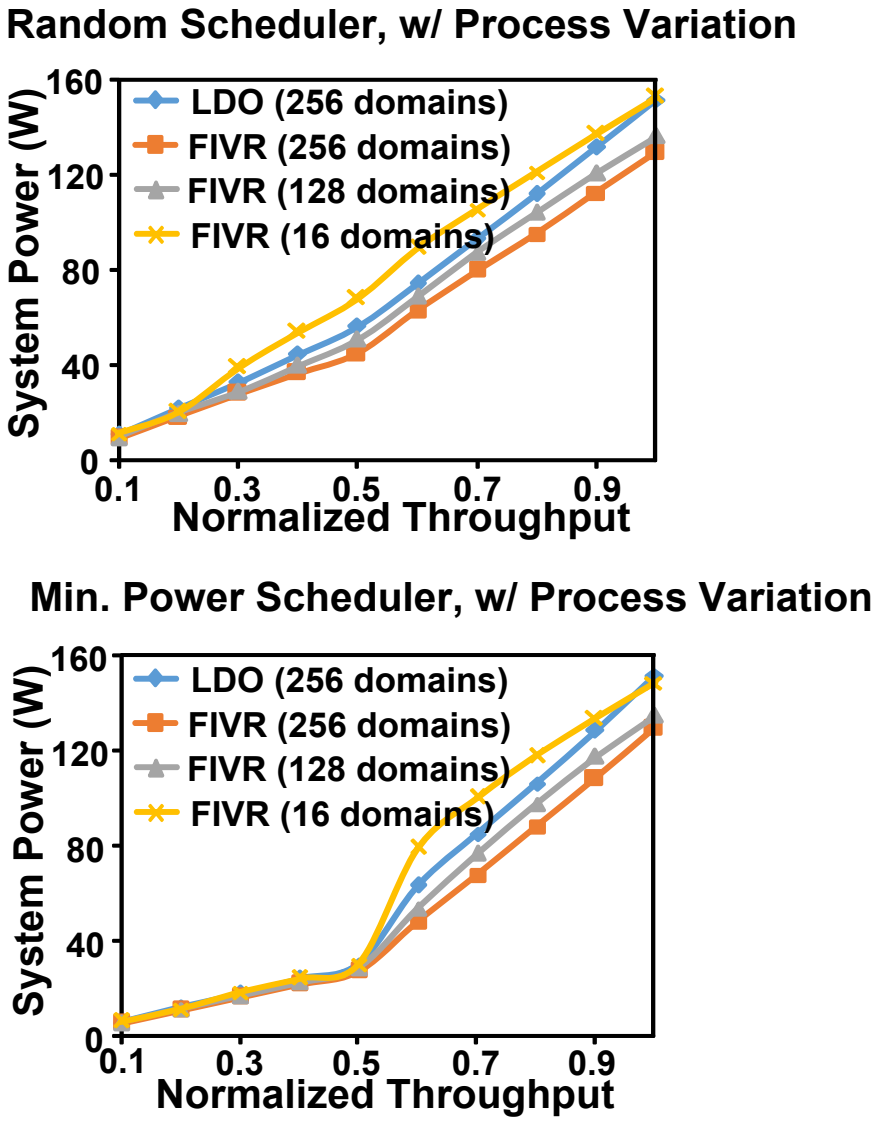


Figure 5.17: System power vs. throughput for various power delivery architectures (w/ process variation).

In this work, we compare a FIVR based power delivery solution with an LDO based one, in terms of system power consumption of a 256-core processor. Our analysis shows that under random scheduling and process variation, for a normalized throughput of 0.5, FIVR based architecture with 2-core per voltage domain burns 10% less power than LDO based architecture with 1-core

per voltage domain. We also conclude that if more power states are available, because of better

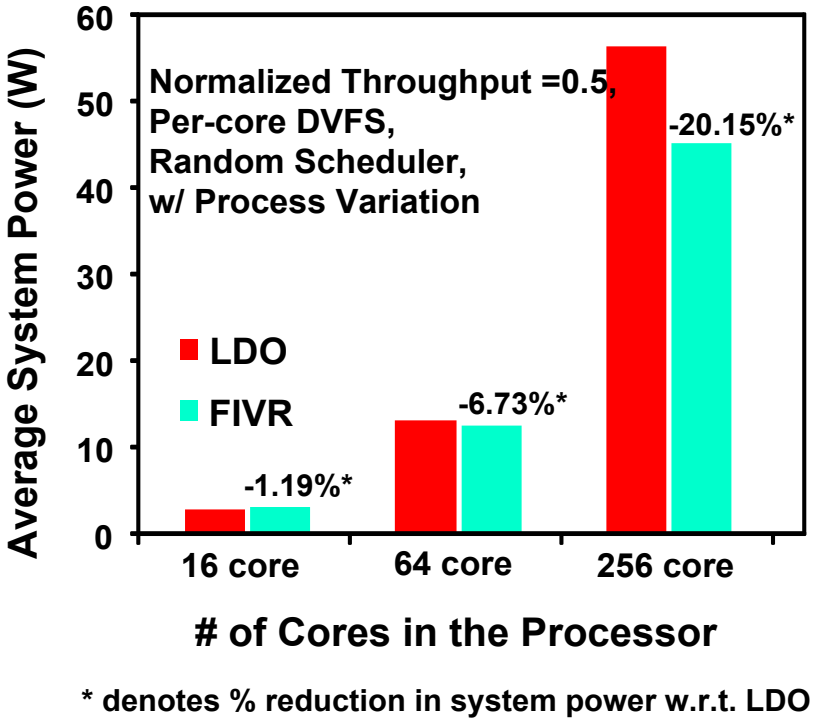


Figure 5.18: Average system power vs. # of cores for per-core LDO and FIVR.

efficiency of FIVR, FIVR architecture with even larger number of cores per domain should be able to outperform per-core LDO architecture.

Although our estimation methodology is sufficient to gain insight into the overall benefits of FIVR and LDO, it could be refined to study specific aspects of the two power delivery methods. For instance, a more detailed power distribution network including on-chip parasitic can be used to incorporate the impact of local supply noise at the expense of further complexity of the model. Another future direction would be to capture the effect of system transient on instantaneous system power in a power budget constrained processor. Finally, since these power-performance characteristics are dependent on the type of scheduler, the analysis should be carried out with application-specific scheduler.

6. CONCLUSIONS

In this thesis, we have presented various improvement aspects of power delivery networks in microprocessors, starting from the reduction of noise from power delivery network to the designing of novel capacitor based switched capacitor DC/DC converters. We have studied the source of resonant supply noise and proposed a method based on staggering activation of cores to mitigate first-droop noise in chapter 2. From our 65nm test-chip, using our proposed technique, we have measured an improvement of 12.7% in first-droop noise in a 2-core processor.

The design of switched capacitor DC/DC converters, which use deep-trench capacitor as the flying capacitor of the converters, has been detailed in chapter 3 and chapter 4. We have proposed a fast DVFS technique using a switched capacitor step-down converter and a technique to reduce IR noise from power supply using a switched capacitor step-up converter. Measurement results from a 32nm test chip shows a 5x improvement in DVFS rise transient and a 45% reduction in IR noise with the help of our proposed techniques. We have gone further and proposed a new architecture for switched capacitor DC/DC converter. Our proposed architecture has single input dual output, and the converter supplies currents to both its outputs with the help of TDM. This converter is capable of delivering twice the output current to one of its outputs than an equivalent SISO converter, when the other output of our proposed converter remains idle. From a 32nm test chip, we have measured a 29% improvement in IR noise using our SIDO TDDM converter because of the distributed nature of the converter.

Finally, in chapter 5 of thesis, we have studied and compared two state-of-the-art on-chip power delivery architectures under discrete DVFS scenario. In a 256-core processor, a FIVR based 2-core per power domain architecture is found to burn 10% less power than an LDO based per-core DVFS architecture at a normalized throughput of 0.5.

REFERENCE

[1] Q. Wu, M. Pedram, X. Wu, "Clock-gating and its application to low power design of sequential circuits," *IEEE Transactions on Circuits and Systems I*, vol. 47, no. 3, pp. 415–420, Mar. 2000.

[2] Kaushik Roy, S. Mukhopadhyay, H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep sub-micrometer CMOS circuits," *Proceedings of the IEEE*, pp 305-327, Feb. 2003.

[3] P. Macken, M. Degrauwe, M. V. Paemel, and H. Oguey. "A voltage reduction technique for digital systems," *IEEE International Solid State Circuits Conference*, pages 238–239, Feb. 1990.

[4] M. Saint-Laurent and M. Swaminathan, "Impact of Power-Supply Noise on Timing in High Frequency Microprocessors," *IEEE Transactions on Advanced Packaging*, vol. 27, no. 1, pp. 135-144, 2004.

[5] J. M. Rabaey, A. Chandrakasan and B. Nikolic, "Digital Integrated Circuits A Design Perspective," 2003.

[6] 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 Systems*, vol. 10, no. 3, pp. 319-326, 2002.

[7] J. Xu, P. Hazucha, M. Huang, et al., "On-Die Supply-Resonance Suppression Using Band Limited Active Damping," *IEEE International Solid-State Circuits Conference*, pp.2238-2245, 2007.

[8] J. Gu, R. Harjani and C. Kim, "Distributed Active Decoupling Capacitors for On-Chip Supply Noise Cancellation in Digital VLSI Circuits," *Symposium on VLSI Circuits*, pp. 216-217, 2006.

[9] N. Kurd, P. Mosalikanti, M. Neidengard, J. Douglas and R. Kumar, "Next generation Intel® core™ micro-architecture (Nehalem) clocking," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1121-1129, Apr. 2009.

[10] J. Gu, H. Eom and C.H. Kim, "On-chip Supply Noise Regulation Using a Low Power Digital Switched Decoupling Capacitor Circuit," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 6, pp. 1765-1775, Jun. 2009.

[11] E. Hailu, D. Boerstler, K. Miki, J. Qi, M. Wang and M. Riley, "A circuit for Reducing Large Transient Current Effects on Processor Power Grids," *IEEE International Solid-State Circuits Conference*, Feb. 2006, pp. 2238-2245.

[12] M. Ang, R. Salem, and A. Taylor, "An on-chip voltage regulator using switched decoupling capacitors," *IEEE International Solid-State Circuits Conference*, pp. 438–439, Feb. 2000.

[13] K. L. Wong, T. Rahal-Arabi, M. Ma, and T. Greg, "Enhancing microprocessor immunity to power supply noise with clock/data compensation," *Journal of Solid-State Circuits*, vol. 41, no. 4, pp. 749–758, Apr. 2006.

[14] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: a framework for architectural-level power analysis and optimizations," *International Symposium on Computer Architecture*, pp. 83-94, Jun. 2000.

[15] A. J. KleinOsowski, and D. J. Lilja, "MinneSPEC: A new SPEC benchmark workload for simulation-based computer architecture research," *Computer Architecture Letters*, vol. 1, pp. 7–7, 2002.

[16] SPEC 2000: http://www.spec.org/cpu2000/

[17] L. Chang, R. K. Montoye, B. L. Ji, A. J. Weger, K. G. Stawiasz, et al., "A Fully-Integrated Switched-Capacitor 2:1 Voltage Converter with Regulation Capability and 90% Efficiency at 2.3A/mm^2," *Symposium on VLSI Circuits*, pp. 55-56, 2010.

[18] B. Kim, W. Xu, and C. Kim, "A 32nm, 0.9V Supply-Noise Sensitivity Tracking PLL for Improved Clock Data Compensation Featuring a Deep Trench Capacitor based Loop Filter," *Symposium on VLSI Circuits*, 2013.

[19] C. Pei, R. Booth, H. Ho, N. Kusaba, L. Xi, et al., "A Novel, Low-cost Deep Trench Decoupling Capacitor for High-performance, Low-power Bulk CMOS Applications," *International Conference on Solid-State and Integrated-Circuit Technology*, pp.1146-1149, 2008.

[20] N. Butt, K. McStay, A. Cestero, H. Ho, W. Kong, et al., "A 0.039um^2 High Performance eDRAM Cell based on 32nm High-K/Metal SOI Technology," *International Electron Devices Meeting*, pp. 27.5.1-4, 2010.

[21] B. Jayaraman, S. Gupta, Y. Zhang, P. Goyal, H. Ho, et al., "Performance Analysis and Modeling of Deep Trench Decoupling Capacitor for 32 nm High-performance SOI Processors and Beyond," *International Conference on IC Design & Technology,* pp.1-4, 2012.

[22] P. Favrat, P. Deval, and M. J. Declercq, "A High-Efficiency CMOS Voltage Doubler," *IEEE Journal of Solid-State Circuits,* vol. 33, no. 3, pp. 410-416, Mar. 1998.

[23] H.-P. Le, M. Seeman, S. R. Sanders, V. Sathe, S. Naffziger, et al., "A 32nm Fully Integrated Reconfigurable Switched-Capacitor DC-DC Converter Delivering 0.55W/mm^2 at 81% Efficiency," *IEEE International Solid-State Circuits Conference*, pp. 210-211, 2010.

[24] D. Somasekhar, B. Srinivasan, G. Pandya, F. Hamzaoglu, M. Khellah, et al., "Multi-Phase 1 GHz Voltage Doubler Charge Pump in 32 nm Logic Process", *IEEE Journal of Solid-State Circuits,* vol. 45, no. 4, pp. 751-758, 2010.

[25] D. El-Damak, S. Bandyopadhyay, A. P. Chandrakasan, "A 93% efficiency reconfigurable switched-capacitor DC-DC converter using on-chip ferroelectric capacitors," *IEEE International Solid-State Circuits Conference*, pp. 374-375, Feb. 2013.

[26] C.-W. Kuan, and H.-C. Lin, "Near-independently regulated 5-output single-inductor DC-DC buck converter delivering 1.2W/mm^2 in 65nm CMOS," *IEEE International Solid-State Circuits Conference*, pp. 274-275, Feb. 2012.

[27] D. Lu, Y. Qian, Z. Hong, "An 87%-peak-efficiency DVS-capable single-inductor 4-output DC-DC buck converter with ripple-based adaptive off-time control," *IEEE International Solid-State Circuits Conference*, pp.82-83, Feb. 2014.

[28] J. Lee, G. Hatcher, L. Vandenberghe, et al., "Evaluation of Fully-Integrated Switching Regulators for CMOS Process Technologies," *IEEE Transactions on Very Large Scale Integration Systems*, vol.15, no.9, pp.1017-1027, Sep. 2007.

[29] E. A. Burton, G. Schrom, F. Paillet, et al., "FIVR — Fully integrated voltage regulators on 4th generation Intel® Core™ SoCs," *Applied Power Electronics Conference*, pp.432-439, Mar. 2014.

[30] E. J. Fluhr, J. Friedrich, D. Dreps, et al., "POWER8™: A 12-core server-class processor in 22nm SOI with 7.6Tb/s off-chip bandwidth," *IEEE International Solid-State Circuits Conference*, pp.96-97, Feb. 2014.

[31] A. A. Sinkar, H. Wang, N. S. Kim, "Workload-aware voltage regulator optimization for power efficient multi-core processors," *Design Automation and Test in Europe*, pp.1134-1137, Mar. 2012.

[32] W. Lee, Y. Wang, M. Pedram, "VRCon: Dynamic reconfiguration of voltage regulators in a multicore platform," *Design Automation and Test in Europe*, Mar. 2014.

[33] W. Kim, M. Gupta, G.-Y. Wei, et al., "System Level Analysis of Fast, Per-Core DVFS using On-Chip Switching Regulators," *High Performance Computer Architecture*, pp. 123-124, Feb. 2008.

[34] N. Kurd, M. Chowdhury, E. Burton, et al., "Haswell: A Family of IA 22nm Processors", *IEEE International Solid-State Circuits Conference*, pp. 112-113, Feb. 2014.

[35] K. Yao, M. Ye, M. Xu, et al., "Tapped-inductor buck converter for high-step-down DC-DC conversion," *IEEE Transactions on Power Electronics*, vol.20, no.4, pp.775-780, Jul. 2005.

[36] G. A. Rincon-Mora, "Current efficient, low voltage, low drop-out regulators," *Dissertation*, School of Electrical and Computer Engineering, Georgia Institute of Technology, 1996.

[37] R. Sheehan. Nov. 1999, "Active Voltage Positioning Reduces Output Capacitors," Linear Technology, www.linear.com/docs/5600.

[38] NVIDIA Corp., "Tesla Server Solutions", http://www.nvidia.com/object/tesla-servers.html

[39] Intel Corp., "Intel® Xeon Phi™ Product Family: Product Brief",http://www.intel.com/content/www/us/en/high-performance-computing/high-performance-xeon-phi-coprocessor-brief.html

[40] G. Yan, Y. Li, Y. Han, et al., "AgileRegulator: A hybrid voltage regulator scheme redeeming dark silicon for power efficiency in a multicore architecture," *High Performance Computer Architecture*, Feb. 2012.

[41] K. J. Kuhn, M. D. Giles, D. Becher, et al., "Process Technology Variation," *IEEE Transactions of Electron Devices*, vol.58, no.8, pp.2197-2208, Aug. 2011.

[42] S. Pant and E. Chiprout, "Power Grid Physics and Implications for CAD," *Design Automation Conference*, pp. 199-204, Jul. 2007.

[43] S. S. Kudva, R. Harjani, "Fully-Integrated On-Chip DC-DC Converter With a 450X Output Range," *IEEE Journal of Solid-State Circuits*, vol.46, no.8, pp.1940-1951, Aug. 2011.

[44] H.-P. Le, J. Crossley, S.R. Sanders, and E. Alon, "A sub-ns response fully integrated battery-connected switched-capacitor voltage regulator delivering 0.19W/mm^2 at 73% efficiency," *IEEE International Solid-State Circuits Conference*, pp. 372-373, Feb. 2013.

[45] S. Pant, "Design and Analysis of Power Distribution Networks in VLSI Circuits," *Dissertation,* Electrical Engineering, University of Michigan, Ann Arbor, 2008.