High Efficiency Power Delivery for Chip Multiprocessors Using Voltage Stacking

High Efficiency Power Delivery for Chip Multiprocessors using Voltage Stacking

A dissertation presented

by

Sae Kyu Lee

to

John A. Paulson School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University

Cambridge, Massachusetts

July 2016

Thesis advisor

Gu-Yeon Wei and David Brooks

Author

Sae Kyu Lee

High Efficiency Power Delivery for Chip Multiprocessors using Voltage Stacking

Abstract

Efficient power delivery is a critical design target for modern computing systems. Inefficiencies in the power delivery network, coupled with increases in current draw and worsening current transients, significantly hinder energy-efficient power delivery for modern computing systems. Furthermore, off-chip components ostensibly have not scaled, exacerbating the power delivery challenges facing future computing systems.

This thesis explores *voltage stacking*, an alternative power delivery solution that stacks voltage domains in series to deliver a high voltage to the chip rather than delivering a low voltage to the voltage domains in parallel. This results in reduction of overall current draw and alleviates off-chip power delivery inefficiencies. However, the trade-off of voltage stacking is increased complexity in mitigating inter-layer voltage noise and in efficient inter-layer communication.

This thesis starts by exploring the characteristics of inter-layer voltage noise in voltage stacking. Results from a test chip prototype demonstrate the inherent properties of voltage stacking and show that supply rail impedance in voltage stacked systems depend on the overall current flow through the stack, leading to better noise immunity for high-throughput workload scenarios with higher overall power due to lower impedance across the voltage stack.

Based on this understanding, the thesis presents a 16-core test chip that utilizes adaptive frequency clocking scheme with an efficient switched-capacitor DC-DC

converter to mitigate voltage noise and improve system performance and efficiency. Experimental results demonstrate robust voltage noise mitigation and better than 94% power delivery efficiency for high-throughput workload scenarios. Results also illustrate that augmenting the hardware techniques with intelligent workload allocation can preemptively reduce the inter-layer activity mismatch and further improve system power delivery efficiency.

Next, to address the issue of efficient inter-layer communication, the thesis presents coupled inter-layer communication schemes. A test chip prototype demonstrates that by transmitting pulsed signals across capacitors and inductors to isolate the stacked voltage domains, the coupled layer shifters reliably shift signals across stack layer boundaries and enable efficient broadcasting of signals from one to multiple stack layers.

Finally, the thesis concludes with future directions that can further improve performance and on-chip power delivery efficiency of voltage stacked systems.

Contents

List of Figures

List of Tables

Acknowledgments

Looking back on my PhD, I realize how incredibly fortunate I was to have had this opportunity, but more importantly how lucky I was to have shared this journey with so many wonderful people. It would not have been possible to complete this journey that I unknowingly undertook seven years ago without the people around me.

I am indebted to my advisors Gu and David for their guidance and advice through my PhD. They were always willing to provide me with invaluable feedback on my research and analysis and to engage and challenge me to grow as a researcher. I grew so much under their guidance during my time here, and it inspires me that their passion and enthusiasm for research hasn't waned one bit over the seven years I've been here. Seven years of PhD certainly doesn't come without ups and a fair share of downs, and I thank Gu and David for their guidance and patience over the years to see this through til the end.

I would also like to thank professor Hanspeter Pfister for being on both my qualifying exam committee and my dissertation committee. He was kind enough to see me through the two important milestones during my PhD and provided me with valuable feedback on the thesis.

I can't thank enough all my lab mates of past and present, for making the research lab such a fun and vibrant environment to work in. I learned so much from the diverse expertise offered by my lab mates, but more importantly they enriched my lab life and made the journey so enjoyable. Where ever I end up later on, I will miss the daily group lunches, the whiskey Fridays, impromptu group meetings at a nearby bar and the daily conversations chatting about anything and everything. I was especially fortunate to have shared the bulk of my time here with Wonyoung, Tao, Mario and

Silvia, with whom I shared countless technical discussions and many a sleep deprived tape-out nights. I feel especially fortunate to have worked with Tao on my final voltage stacking chip, without which this thesis would not have been possible.

I would like give special thanks to Glenn for all his help and computer wizardry during my PhD. I don't know what I (or our lab in general) would have done without Glenn's help in server and tool setup and trouble shooting. I can't remember how many time I had to urgently call him to try and fix some problem I was having with the tools or the servers, and he would always come through, no matter what time of day it was. I'd like to also thank him for his patience through all my dumb linux and tool questions at all hours of the day.

I am also truly blessed to have met so many wonderful friends outside of the lab, who have been like a second family to me. Jaehyoung, Wonbin, Bumjin, Jaeyoung and Honggeun, who I was fortunate enough to meet in my first year, Taeyeon, Changwook, Sunah, Eunsil, Rosa, who I met in subsequent years. I have lasting memories of our drunken late nights on weekends and all the fun we had during my time here. I could always count on them to be there for me through all the highs and the lows of the PhD and I will forever cherish their friendship.

I would like to thank my family back in Korea, for all their unwavering love and support not only throughout my PhD, but throughout my life up to this point. My parents, my brother and sister, who always believed in me no matter what, and my only regret is that I have not been able to go back to visit them more often during my PhD. I'd also like to thank my extended family, my brother-in-law Eric for his kindness and support, and my niece Sena, whose pictures and videos put a smile on

my face at any time, and my mother-in-law for her support and advice to push me through the last stretch.

Finally, I'd like to thank my wife Yuhree who I met during my 4th year in graduate school. I can't imagine how I would have got through the last half of my PhD without her in my life. Her love and support embolden me everyday, and I am truly blessed to be able to embark on the next chapter of my life with her. As I close this chapter of my life, no matter what comes of this, my decision to pursue a PhD was the best decision I ever made because it led me to her.

Dedicated to Yuhree, my muse.

Chapter 1

Introduction

For years, in addition to architecture and circuit-level innovations, the computing industry has relied heavily on process and clock frequency scaling for consistent performance improvements. However, at the turn of the century, traditional Dennard scaling slowed and increases in clock frequency incurred excessive power penalties. Coupled with the thermal limits of air cooling, which limits the maximum power dissipation of microprocessors, these factors and the inevitable rise in chip power consumption resulted in the computing industry hitting the proverbial "Power Wall". As a result, clock frequency scaling all but ceased and the maximum power consumption of modern high performance microprocessors have stayed relatively constant ever since. As Moore's Law struggles on to provide designers with the capability to integrate more transistors and functionality on a single chip, the computing industry is now forced to shift focus away from improving single-threaded performance to exploiting higher levels of parallelism so that progress may continue in this power-constrained environment. The idea is that if workloads can be broken

up into smaller pieces, chip multiprocessors can operate in parallel, offering higher aggregate throughput while consuming less energy per piece. Thus, the industry has been pushing towards many-core chip multiprocessors, integrating more and more cores with the additional transistors afforded every process generation [6, 14].

However, even with this shift in philosophy, the continued rise in power density due to device scaling and integration threatens to limit the amount simultaneously active cores on die [16]. Therefore, to keep power at manageable levels within fixed power budgets, there is growing interest in aggressive supply voltage scaling beyond the traditional supply voltage scaling provided by process technology. Such scaling looks to operate the transistors in the sub-threshold or near-threshold regime, the idea being that lost performance due to voltage scaling can be reclaimed by an increase in aggregate throughput. In addition, modern processors also employ a myriad of aggressive power-saving techniques, such as power-gating and clock-gating to reduce unnecessary waste in power and energy. These trends have led to an overall increase in the average current draw of the chip and to worsening current transients in the power-constrained environment of modern processors.

As a result, efficiently delivering power to processors has become a critical design target due to the inefficiencies in the power delivery network, which forces stringent requirements on the power delivery impedance to keep supply noise and power losses within budget. Exacerbating the issue, the off-chip components do not scale, conflicting with the the ever-decreasing power delivery impedance requirements needed to keep up with the high current demands of modern computing systems.

Voltage stacking is an alternative on-chip power delivery solution that stacks volt-

age domains in series to provide a single high voltage to the chip. By recycling charge through the series-connected stack layers, voltage stacking offers several advantages. For the same chip power, an n-way stacked system reduces the current draw of the chip proportionally by n, reducing off-chip losses and alleviating power delivery impedance requirements. It also obviates a high step-down off-chip DC-DC converter. For high-power systems that draw power from a high-voltage rail (e.g., 12V), lower step-down ratios can improve off-chip regulator efficiency. For battery-powered applications, there is an opportunity to power the chip directly off of the battery, eliminating all together losses caused by off-chip voltage conversion and saving precious board space. However, voltage stacking presents two significant challenges in its application to future chip multiprocessors: i) internal voltage noise due to inter-layer activity mismatch, and ii) inter-layer communication.

Voltage noise is a big concern facing modern computing systems. Typically in synchronous digital systems, the maximum operating frequency of the system is determined by the minimum supply voltage (V_{MIN}) observed in the system. This is because correct functionality must be guaranteed under even the worst-case voltage noise conditions. Therefore, processors typically operate with a sufficient voltage margin to account for the worst-case voltage noise of the system, which leads to wasted energy and loss in the overall energy-efficiency. Traditionally, in conventional digital systems, voltage noise is caused by the interaction between the current driving through the off-chip power delivery network, and the parasitic inductance and resistance of the components along the power delivery path. Voltage stacking alleviates these conventional voltage noise issues by reducing the magnitude of the load current.

However, due to its series-connected nature, voltage stacked systems are susceptible to internal voltage noise, caused by inter-layer activity mismatch, which has inherently different attributes to voltage noise in conventional systems. Therefore, it is critical to understand the attributes of internal voltage noise in voltage stacking and to mitigate voltage noise using efficient hardware techniques.

Furthermore, efficient communication between the cores, and between cores and the memory subsystem is critical to maintain cache coherency and to maximize thread-level parallelism in modern high-performance chip multiprocessors. With continued integration of more cores on a single chip, the communication requirements could soon become a bottleneck for extracting thread-level parallelism. Voltage stacking adds an extra level of complexity to on-chip communication because of the fact that cores in different stack layers operate at different voltage reference planes. Therefore, efficient on-chip signaling for inter-layer communication is required for the adoption of voltage stacking in future many-core systems.

This thesis explores voltage stacking as an efficient power delivery solution and addresses these key issues associated with voltage stacking. The thesis presents hardware techniques to tackle the challenges of internal voltage noise mitigation and inter-layer communication to enable the application of voltage stacking in future chip multiprocessors.

1.1 Organization

Chapter 2 presents background on the inefficiencies and challenges related to power delivery in modern chip multiprocessors, and describes the significant benefits that

can be achieved by voltage stacking. The chapter then reviews the main challenges associated with voltage stacking, namely voltage noise due to inter-layer activity mismatch, and inter-layer communication. The chapter concludes by presenting a thorough literature review of the prior arts in the area of voltage stacking.

Chapter 3 explores the key attributes of voltage noise in voltage stacking based on a test chip prototype implementing a 3-layer voltage stacked system. The results illustrate the basic properties and dependencies of internal voltage noise in voltage stacking, and shows that unlike conventional power delivery schemes, supply rail impedance in voltage stacked systems depend on aggregate current flow through the voltage stack, leading to better noise immunity for high-throughput (or high-power) workload scenarios for many-core processors.

Chapter 4 presents the application of an adaptive frequency clocking scheme with an efficient switched-capacitor (SC) DC-DC converter to mitigate noise on the stack layers and to improve system performance and efficiency in the context of high-throughput multicore systems. Experimental results from a 16-core test chip demonstrate robust voltage noise mitigation and showcases voltage stacking as a high efficiency power delivery scheme. By efficiently recycling charge through the stack layers, the system achieves peak efficiency of >99% and achieves better than 94% power delivery efficiency even when SC converter efficiency is limited to only 68%. This chapter also illustrates that augmenting the hardware techniques with intelligent workload allocation, which exploits the inherent properties of voltage stacking, can preemptively reduce inter-layer activity mismatch and further improve system efficiency.

Chapter 5 describes pulse-mode coupled communication techniques to address the issue of efficient inter-layer communication in voltage stacked systems. The layer shifting communication circuits transmit pulsed signals across capacitors and inductors to isolate stacked voltage domains and reliably shift signals across stack layer boundaries. These schemes also enable efficient broadcasting of signals from one to multiple stack layers with minimal overhead. Measurement results from a test chip prototype implementing a 3-layer voltage-stacked inter-layer communication system demonstrate efficient inter-layer communication from one to three layers at 3.4Gb/s.

Finally, Chapter 6 concludes with future directions that can further improve voltage noise mitigation and power delivery efficiency of voltage stacked systems to fully realize voltage stacking as high efficiency power delivery solution for future chip multiprocessors.

Chapter 2

Overcoming the Challenges facing Microprocessor Power Delivery using Voltage Stacking

For decades, the semiconductor industry reaped the benefits of process scaling that provided simultaneously faster and lower power transistors. While transistors continue to scale smaller, traditional Dennard scaling, which was at the heart of process scaling of previous years, has all but ended. Furthermore, while traditional supply voltage scaling in accordance with Dennard scaling has also stopped, voltage scaling in process technology continues, albeit at a slower rate, spurred by the progress in the advanced technology nodes. ITRS projections for supply voltage scaling of process technology, which are presented in Figure 2.1 [1] and which closely track actual industry trends, show that supply voltage will continue to scale down to 0.65V by year 2022. As supply voltage continues to scale, the overall current draw of a

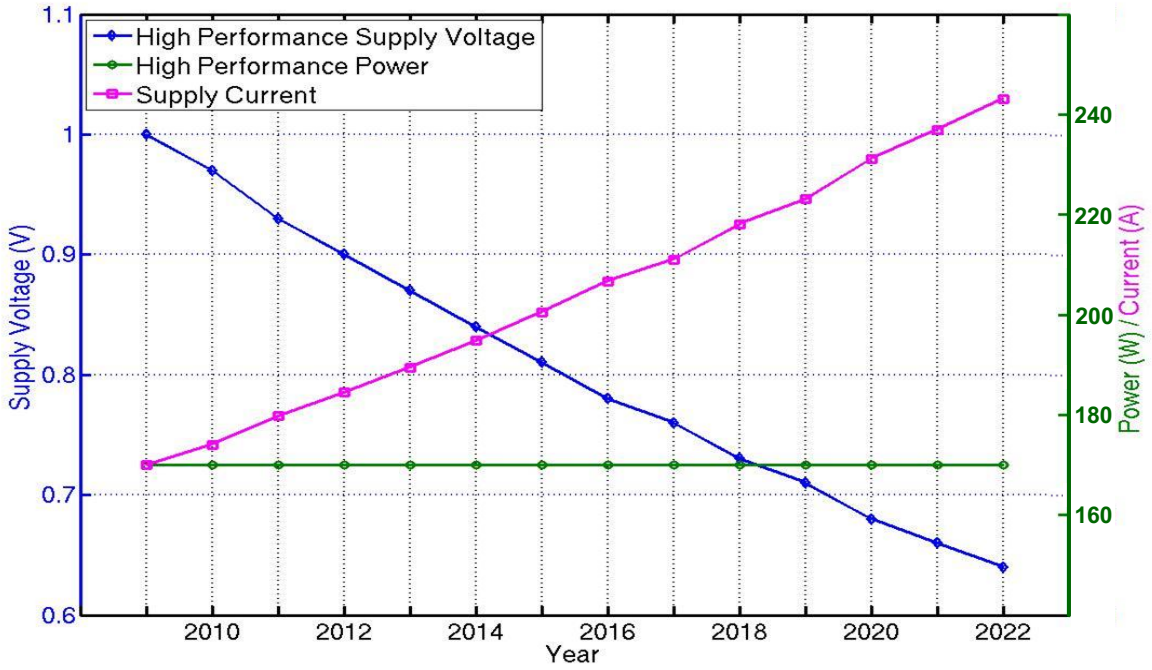


Figure 2.1: 2009 ITRS projection of supply voltage scaling, and project current draw of 165W processor [1].

chip increases for a fixed power budget. As shown in Figure 2.1, a high-performance microprocessor with maximum power consumption of 165W currently draws in excess of 200A and is projected to draw more than 240A in 2022. Furthermore, there is growing interest in scaling the voltage even further to operate transistors more energy-efficiently in the near-threshold or sub-threshold regime. The idea behind this is that lost performance due to voltage scaling can be reclaimed through an increase in aggregate throughput via a higher number of simultaneously active cores [15, 27, 44]. Therefore, the maximum current draw of future high-performance microprocessors projects to be even larger than that shown in Figure 2.1.

In addition to voltage scaling, modern processors employ aggressive power management techniques to reduce power consumption and improve energy efficiency. Dy-

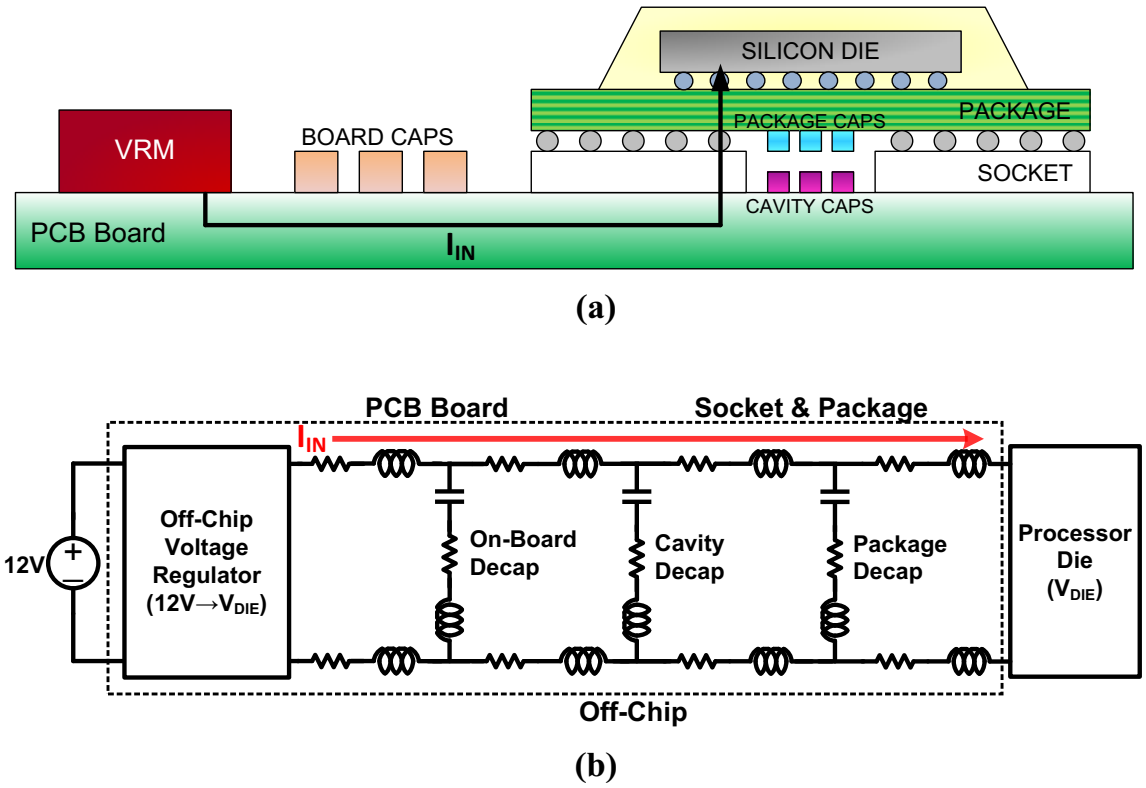


Figure 2.2: Block diagram of (a) microprocessor power delivery network and (b) schematic illustrating parasitics of the power delivery path.

namic voltage and frequency scaling (DVFS) adapts both the voltage and frequency of the system with respect to changes in workloads [68, 30]. Furthermore, clock gating and power gating mitigate unnecessary switching and leakage power of idle logic blocks. While these power management techniques can drastically reduce the power consumption of a system, they increase fluctuations in the current consumption of the chip because of changes in processor activity.

The continued increase in overall current consumption, coupled with increases in current fluctuations, present significant challenges to efficient power delivery. These challenges are further magnified because of the magnitude of a current that must

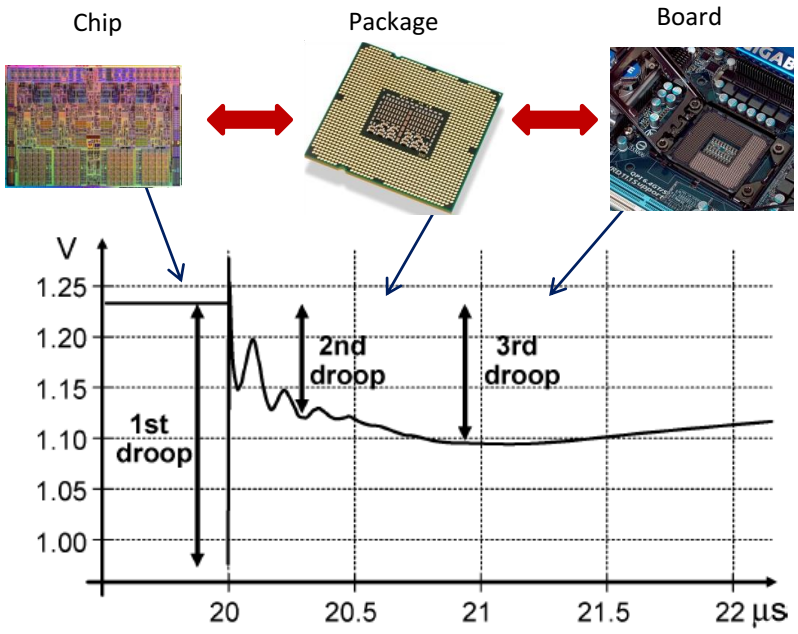


Figure 2.3: Waveform illustrating *di/dt* noise.

be supplied to the chip through the off-chip power delivery network. This chapter provides a detailed explanation of the inefficiencies associated with power delivery in modern microprocessors and how these inefficiencies threaten to drain performance and power gains of modern chip multiprocessors (Section 2.1). This chapter then discusses how voltage stacking can provide a scalable solution to alleviate the identified issues and to provide a path for continued gains in performance and power (Section 2.2). The chapter concludes with a thorough review of the prior arts in voltage stacking (Section 2.4).

2.1 Inefficiencies of Delivering Power to Modern Microprocessors

To fully understand the challenges associated with power delivery, Figure 2.2(a) presents a block diagram illustration of a conventional microprocessor power delivery network. For desktop and server type systems, power is converted from the wall plugs (typically 110VAC or higher for datacenters), to 12V DC and is delivered to the off-chip voltage regulator module (VRM). This conversion stage is not shown in Figure 2.2 for sake of brevity. The VRM then steps down the 12V supply to the voltage required by the chip. Then the VRM delivers the power via the PCB board, socket, package, and finally through the Controlled Collapse Chip Connection (C4) bumps to the silicon die.

These off-chip components along the power delivery path have parasitic resistances and inductances, as shown in the schematic representation presented in Figure 2.2(b). These component parasitics interact with the current driving through the off-chip connections, which in turn causes inefficiencies in the power delivery that reduces the overall system efficiency.

- **IR drop:** As supply current increases, the IR drop due to off-chip parasitic resistance increases. As supply voltage continues to scale, maintaining the IR drop within a manageable margin becomes critical, as it eats into the processors operating voltage. For a 165W processor operating at 0.8V, keeping the worst-case IR drop to within 10% of the operating voltage requires the overall power delivery resistance to be less than 0.4mΩ.

- **I^2R power loss:** In addition to IR drop, a power delivery network incurs I^2R power loss because of off-chip parasitic resistance, where the voltage margin percentage for a set IR drop determines the worst-case I^2R power loss. For a 165W processor, a 10% IR drop margin translates into 10% of maximum power wasted on the power delivery network at full load, which is equivalent to an excess of 16W.

- ***di/dt* noise:** *di/dt* voltage noise is caused by fluctuations in current consumption of the chip interacting with parasitic inductances along the off-chip current delivery path, which causes a voltage drop across the inductors. The magnitude of the *di/dt* noise (ΔV) is determined by:

$$\Delta V \propto \Delta I \cdot \sqrt{L/C} \tag{2.1}$$

  where, ΔI is the magnitude of the change in current, L is the loop inductance of the components in the power delivery networks, and C is the decoupling capacitance. As shown in Figure 2.2(b), to mitigate *di/dt* noise, decoupling capacitors are placed at various stages of the power delivery network to reduce the impedance of the power delivery network across a wide range of target frequencies. These capacitors act as a charge reservoir providing the instantaneous charge required by the processor locally so that the charge does not need to come through the inductance of the power delivery network. As decoupling capacitors at each stage are depleted, the change in current flow through the inductors cause successive voltage drops as shown in the example voltage noise waveform for a step current load presented in Figure 2.3. To ensure correct

functionality, designers must impose a sufficient voltage margin for the worst-case voltage noise, limiting the maximum operating frequency and degrading overall performance and efficiency.

- **Electromigration:** Rising current levels can also cause electromigration (EM), which causes long-term reliability issues because of the degradation of the C4 bumps over its lifetime [67, 12]. To avoid EM issues, more C4 bumps must be allocated for power to counteract the increase in current. However, with the C4 bump density expected to remain relatively constant for the foreseeable future [1], chip designers must carefully balance the increasing I/O demands with the reliability requirements of future microprocessors.

One way to address the inefficiencies related to voltage noise and power loss in power delivery is to force stringent off-chip impedance requirements to satisfy the voltage noise margin and power loss demands of future computing systems. As supply voltage scales, power delivery impedance requires a quadratic reduction to maintain IR drop and *di/dt* noise within the same percentage margin. In addition, reducing I^2R power loss may require further reduction in resistance. However, the off-chip components, such as off-chip decoupling capacitance, PCB trace, package and sockets, has ostensibly not scaled and are projected to remain relatively constant. This exacerbates the power delivery inefficiencies and challenges ahead.

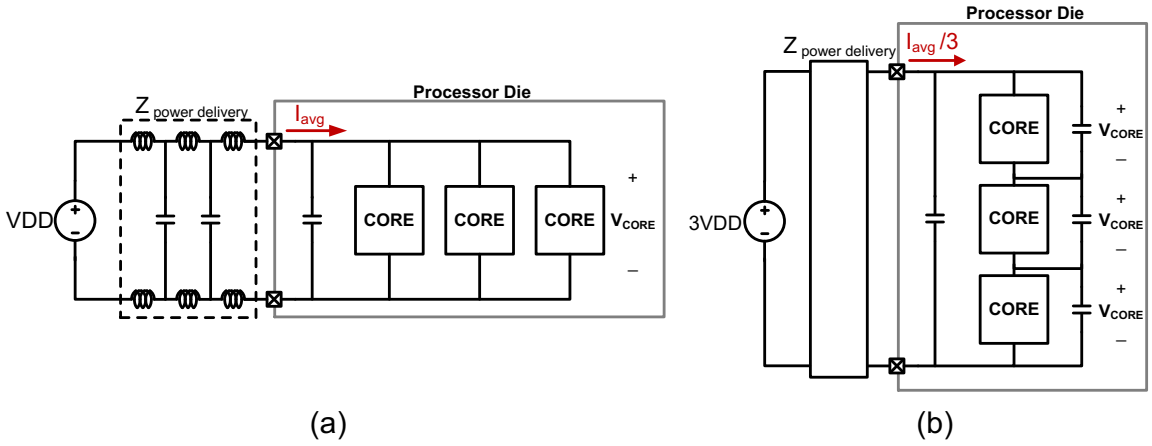


Figure 2.4: Block diagrams of (a) Conventional Power Delivery Scheme, (b) Voltage Stacking

2.2 High-voltage Power Delivery via Voltage Stacking

An alternate method by which to alleviate the inefficiencies in the power delivery network is to reduce the chip current draw that interacts with the parasitics, as opposed to reducing the impedance of the off-chip components. A simple solution to reducing the chip current draw is to supply power to the chip at a higher voltage. This reduces the current draw given the same power consumption.

Voltage stacking is an energy efficient approach to high voltage power delivery which stacks voltage domains in series to provide a single high voltage to the chip instead of delivering power to parallel loads [46, 24, 11, 33, 28, 73, 57, 38, 34, 63, 5]. Figure 2.4 presents a conceptual block diagram of voltage stacking for a three core example. Instead of supplying power to the cores in parallel at VDD, the cores are stacked on top of one another and a three-times higher voltage is applied across the

entire stack. For the same total chip power, the average current flowing through the power-delivery network is reduced by a factor of three. An n-way stacked system would reduce the current draw of the chip proportionally by n. By delivering a higher voltage to the chip and reducing the chip current draw, voltage stacking addresses the key challenges facing power delivery for future microprocessors:

- IR drop across the power delivery subsystem caused by board and package resistance is reduced by a factor of n. In a conventional power delivery scheme, IR drop across the power delivery subsystem affects all cores equally. However, in a voltage stacked design, the effect of the IR drop across the voltage stack is distributed across the n-stack layers. Therefore, IR drop across a core operating voltage (V_{CORE}) is ultimately reduced by a factor of n^2.

- I^2R power loss is also reduced by a factor of n^2 compared to conventional power delivery scheme as seen in Figure 2.4(a), which significantly reduces the power wasted through the power delivery subsystem. For a three-layer voltage stacked system shown in Figure 2.4(a), I^2R power loss reduces by a factor of nine.

- Assuming identical off-chip power delivery network, di/dt noise is reduced by $\sqrt{n}$, despite accounting for reduction in on-chip capacitance by a factor of n caused by the series connected voltage domains.

- Reduction in current alleviates long-term reliability concerns related to electro-migration in C4 bumps.

- Off-chip voltage regulators benefit from lower step-down ratio requirements of voltage stacking, which leads to higher regulator efficiency and reduction in

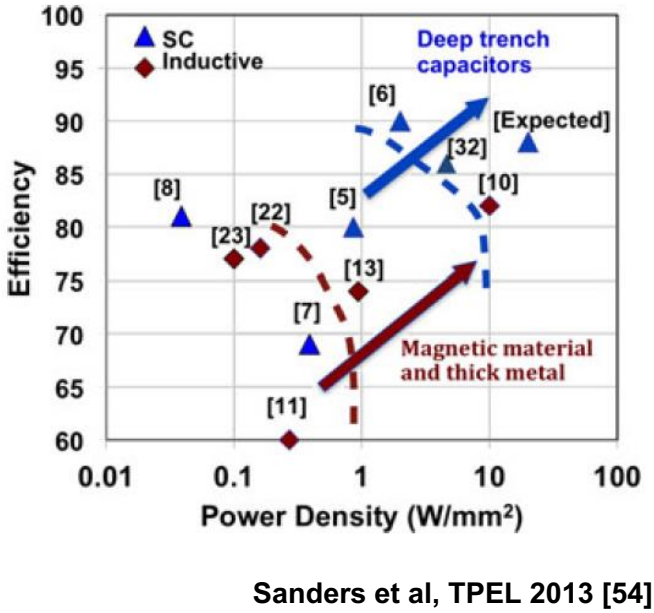


Figure 2.5: Performance of fully integrated voltage regulators in production CMOS and in processes where extra steps allowed.

design complexity. For high-power systems that draw power from a high-voltage rail (e.g., 12V), lower step-down ratios can improve off-chip regulator efficiency. For battery-powered applications, there is an opportunity to power the chip directly off of the battery, eliminating all together the losses caused by off-chip voltage conversion and saving precious board space.

High voltage power delivery, however, requires efficient voltage conversion on-chip, from the high input voltage to the core voltage. There has been growing interest in integrating DC-DC converters on-chip to perform explicit voltage conversions [29, 3, 32, 62, 25, 54]. In this scenario, the overall power delivery efficiency depends wholly on the conversion efficiency of the integrated voltage regulator (IVR), but IVRs usually suffer from inferior conversion efficiency compared to their off-chip counterparts due in part to poor quality inductors and capacitors available on-chip.

Currently, peak conversion efficiency of IVRs are limited to approximately 80%, unless special process technology such as deep trench capacitors or ultra-thick metal for inductors are available. This is illustrated in Figure 2.5. At such low conversion efficiency, the use of on-chip IVRs is problematic as it negates all the power savings achieved by high voltage power delivery.

Voltage stacking, on the other hand, achieves implicit DC-DC conversion by stacking voltage domains and recycling a charge through the stack layers to divide the high input voltage among the series connected voltage domains. This is illustrated in Figure 2.4(b). By obviating the explicit voltage conversion stage, voltage stacking has the potential to achieve high efficiency power delivery unconstrained by the limits of IVR conversion efficiency.

2.3 Challenges of Voltage Stacking

While voltage stacking has the potential to alleviate the inefficiencies related to power delivery, it also introduces additional concerns caused by the series connected power delivery structure. This thesis addresses two main concerns associated with voltage stacked systems.

2.3.1 Voltage Noise due to Inter-layer Activity Mismatch

As mentioned in the previous section, voltage stacking achieves DC-DC conversion by recycling charge through the stack layers and dividing the high input voltage among the series connected voltage domains. Due to the series connected power delivery structure, Kirchhoff's Current Law (KCL) dictates that identical current

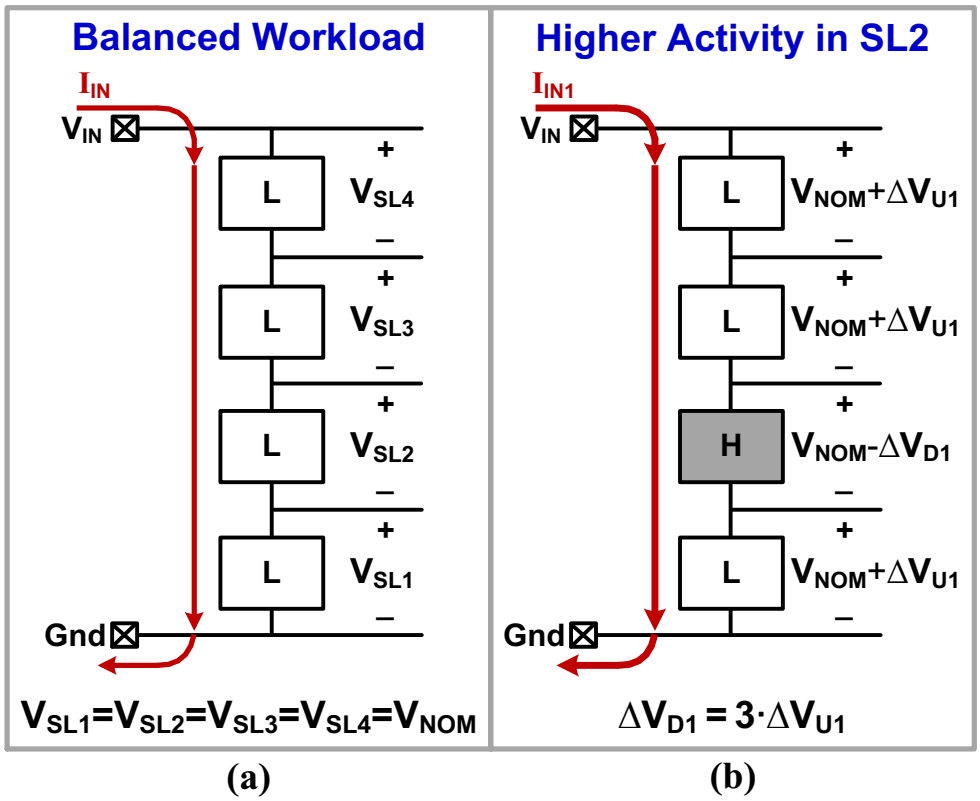


Figure 2.6: Illustration of inter-layer voltage noise in a 4-core 4-way voltage-stacked system using two workload scenarios: (a) Balanced workload across all cores (b) Higher activity in SL2

must flow through all stack layers. Ideally, if the power consumption of all stacked layers perfectly match, the layer voltages evenly subdivide the high input voltage, and voltage stacking achieves optimal implicit step-down conversion with no loss. In practice, inter-layer switching activity mismatch exists and results in fluctuations in the internal nodes that are necessary to maintain identical current through the stack under activity mismatch conditions. Figure 2.6 illustrates inter-layer voltage noise in voltage stacking using a simple 4-core, 4-way voltage stacked system. For simplicity, we assume ideal off-chip connections and assume synchronous digital loads with no leakage current. The system operates off of a fixed V_{IN} of $4\times V_{NOM}$, where V_{NOM} is

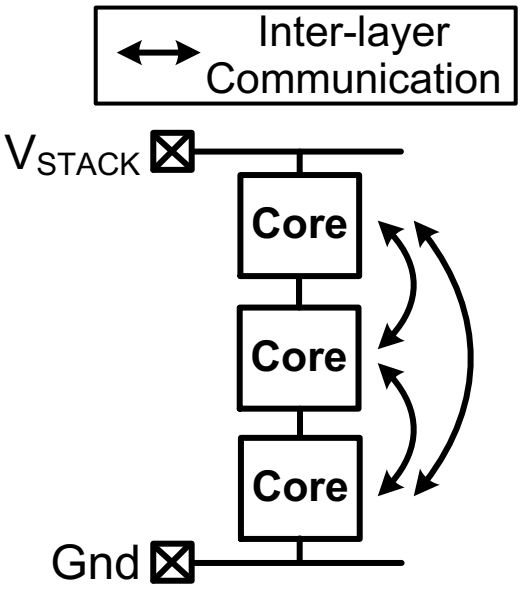


Figure 2.7: Block diagram illustration of the inter-layer communication needs of a 3-core voltage stacked system.

the nominal operating voltage of the cores.

Since identical current must flow through the series-connected cores in the stack, if the power consumption of each stack layer (SL) perfectly matches, the stack layer voltages $V_{SL\{1:4\}}$ will subdivide V_{IN} evenly. This is illustrated in Figure 2.6(a), where all four cores execute the same workload **L** with the same activity factor, α. However, when there is an imbalance in workload activity across the stack layers, layer voltages deviate from V_{NOM} to maintain identical current flow through the stack. Figure 2.6(b) illustrates such a scenario where the core in SL2 executes a higher activity workload **H** with an activity factor of $(\alpha+\Delta\alpha)$. To maintain identical current flow through the stack in this scenario, V_{SL2} decreases, causing voltage droop in SL2, while V_{SL1}, V_{SL3}, and V_{SL4} increase as all stack layer voltages must add up to V_{IN}. With V_{IN} fixed, the droop in V_{SL2} equals the sum of voltage increases in V_{SL1}, V_{SL3}, and V_{SL4}, satisfying

KVL [33]. Note that V_{SL1}, V_{SL3}, and V_{SL4} all increase by the same amount ΔV_{U1} as the cores in these layers all execute the same workload **L**.

This simple example illustrates that any inter-layer activity mismatch manifests itself as voltage noise in the stacked voltage layers. This susceptibility to voltage noise negatively impacts system performance and energy-efficiency, and furthermore, it increases system reliability concerns. This is a critical issue because it can potentially negate any benefits of voltage stacking as an efficient power delivery solution.

2.3.2 Inter-layer Communication

Modern chip multiprocessors require efficient on-chip communication to maximize performance and throughput of the system, whether such communication be for cache-coherency protocols or for core-to-core message passing communication [9, 55, 58, 59, 65, 64, 31, 48, 10, 7]. As process scaling continues to allow for more transistors and greater integration of more cores on a single chip, efficient on-chip communications is critical for future chip multiprocessors. Voltage stacking results in an additional level of complexity for communication between stacked voltage domains that do not share the same reference plane. Figure 2.7 illustrates the communication requirements of a simple 3-core voltage stack as an example. Direct communication between different stack layers requires isolation to avoid DC current paths between stacked voltage rails. Furthermore, direct communication between stacked voltage domains not in close physical proximity requires efficient repeater-less long-distance signaling because repeater placement may not be possible en route.

2.4 Prior Works in Voltage Stacking

With continued supply voltage scaling, voltage stacking has been gaining traction as an alternative on-chip power delivery solution which can alleviate inefficiencies in power delivery caused by increasing current levels. Voltage stacking borrows its core idea from charge recycling. This concept was previously demonstrated in series stacked bus capacitance for ultra high data rate bus architecture [69] and in highly capacitive lines of memories to reduce power dissipation [37, 70].

Rajapandian *et al.* first applied the concept of charge recycling to logic domains in order to achieve implicit DC-DC conversion. They demonstrated voltage stacking within a single 16-by-16 carry-save-array multiplier by partitioning the multiplier and configuring the partitions in a 2-way voltage stacked array [47]. This work was then expanded on to a voltage stacked system consisting of multiple 16-by-16 multipliers with a single multiplier in each stack layer [46]. The latter work demonstrated 2-way and 3-way voltage stacked systems with level shifter circuits that utilized full-swing capacitive coupled inter-layer communication via MOS capacitors, and implemented linear regulators to regulate the internal nodes. However, level shifters based on MOS capacitors suffer from reliability issues caused by device overstress and prohibits direct communication between non-adjacent stack layers. Furthermore, while linear regulators are easy to integrate with low area overhead, they suffer from inherently low efficiency which limits the overall efficiency of the system.

Prior works have focused on various voltage regulation schemes to regulate the internal voltage rails and to overcome the inefficiency of linear regulators. Chang *et al.* [11] demonstrated a 2-to-1 switched-capacitor (SC) converter using deep trench

capacitors and proposed its use in supplying power to two stacked voltage domains. To extend voltage stacking beyond two stacked layers, Zhang *et al.* [73, 74] proposed using multiple 2-to-1 SC converters to support more than two stacked layers at the expense of higher complexity and overhead. On the other hand, Shenoy *et al.* [57] and Kesarwani *et al.* [28] demonstrated off-chip inductor based switching regulators for voltage stacking applications. However, the inherent difficulty of integrating high-quality inductors on-chip hinders full integration of inductor-based regulators.

In addition to voltage regulators, researchers have also explored internal voltage regulation at the system level using software techniques. McClurg *et al.* [38] proposed voltage stacking datacenter servers to remove the per-server conversion losses. To regulate the internal nodes of the stack, the work relied entirely on software techniques to manage the server workloads for purposes of balancing power consumption. Ueda *et al.* [63] demonstrated a battery powered system that used a simple tank cap connected to the intermediate node which would temporarily store energy caused by mismatch current, while also relying on task scheduling to regulate the intermediate voltage of a two stack system. However, relying entirely on software techniques limits the application scope of the system because software techniques cannot guarantee voltage regulation over all possible workload scenarios.

There has also been efforts to extend the application of voltage stacking beyond the traditional on-chip power delivery. Jain *et al.* [24] proposed voltage stacking for delivering power to 3D ICs to reduce voltage noise by voltage stacking 3D ICs and using additional off-chip power supplies to supply the differential power between the stacked chips. Meanwhile, McClurg *et al.* mentioned above, demonstrated voltage

stacked datacenter servers.

This thesis builds upon the prior works on efficient voltage noise mitigation, by not only implementing a fully-integrated voltage regulator but by also exploring system-level techniques including adaptive frequency clocking schemes and intelligent workload allocation schemes to provide a holistic approach to mitigating voltage noise in voltage stacking. In addition, this thesis also explores efficient inter-layer communication schemes to overcome the limitations of level shifter circuits proposed in prior works.

With the understanding of the basics of voltage stacking and its main challenges, we now take a step back to understand the inherent properties of voltage stacking and the characteristics of internal voltage noise.

Chapter 3

Evaluation of Voltage Noise in Voltage-Stacked Multicore Systems

Characterizing and managing voltage noise is crucial for synchronous digital systems, as the clock frequency must be chosen to ensure correct operation under even the worst-case voltage noise conditions. Typically in conventional systems, voltage noise is caused by the interaction between the load current and the parasitic inductance and resistance of the power delivery network. Voltage stacking alleviates these conventional voltage noise issues by reducing the chip current draw. However, series stacked voltage domains are susceptible to voltage noise with inherently different attributes that is orthogonal to conventional voltage noise.

In voltage stacked systems, KCL dictates that identical current must flow through each of the series connected stack layers. However, power consumption of cores can fluctuate as a function of workload, various power-saving schemes (e.g., clock gating) and data, leading to inter-layer activity imbalances. There is an implicit self-

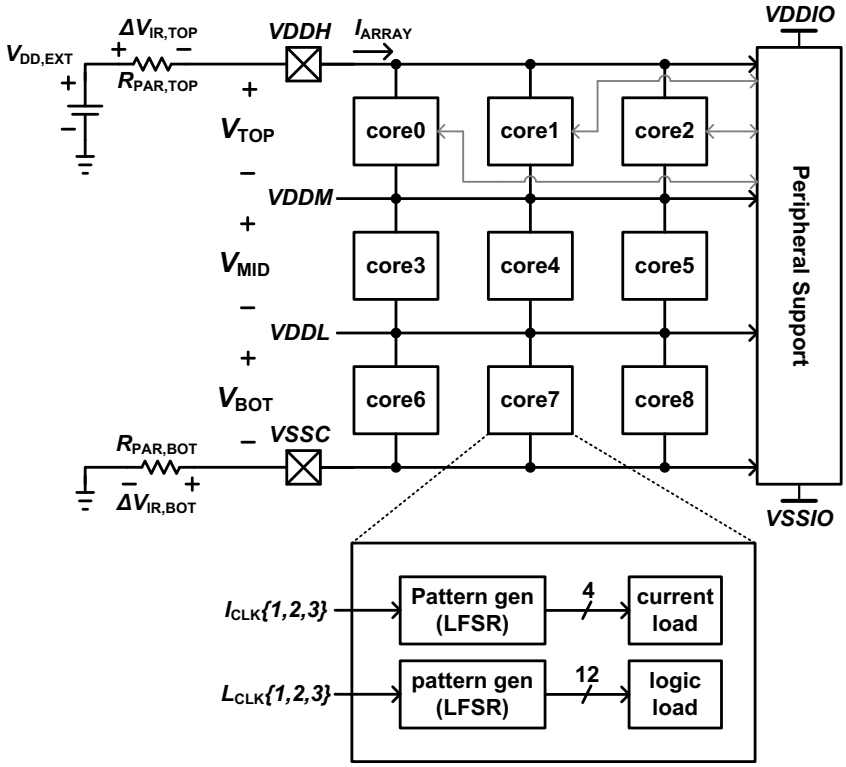


Figure 3.1: Block diagram of the test-chip prototype with a 3x3 voltage stacked array of power-consuming blocks

regulating loop built into the stack that distributes the internal voltage levels to compensate for any activity imbalances between the stack layers. For example, if a stack layer has lower activity compared to other layers, its voltage increases to compensate at the expense of reducing the voltage across the other stack layers. Hence, any aggregate inter-layer core activity mismatch manifests itself as internal voltage noise. Therefore, balancing workload activity across the layers is a key challenge facing voltage stacking.

Given the understanding of the potential problems of voltage stacking, this chapter describes the design of a test-chip protoytpe that evaluates the key attributes and challenges of voltage stacked multicore systems and presents experimental results

from the test chip. The overview of the test chip is presented in Section 3.1, and Section 3.2 presents the evaluation of voltage noise due to static inter-layer activity mismatch. Building upon this voltage noise characterization, Section 3.3 explores the supply rail impedance characteristics of voltage stacked systems that has significant impact on voltage noise.

3.1 Overview of Near-threshold Voltage Stacked Test Chip Prototype

Block diagram of the test chip prototype comprising a 3x3 array of power-consuming *cores* is presented in Figure 3.1. The test chip is organized as three voltage stacked layers with three cores per layer with a total of four voltage rails for the array. Only VDDH and VSSC connect to an external voltage source ($V_{DD,EXT}$) while internal rails VDDM and VDDL settle to levels dictated by the aggregate power consumption of the *cores*. *Peripheral support* circuitry, containing level shifters and voltage monitoring circuits, feeds signals into and reads signals out of the array, and operates off of separate rails VDDIO and VSSIO connected to a separate external source. The test-chip operates at VDDH=1.5V with each *core* nominally operating at 0.5V.

Each *core* contains two sets of power-consuming blocks shown in Figure 3.2. The *logic load* block contains three sets of four independent logic paths, all roughly 15 fanout-of-4 (FO4) inverter delays long. The logic paths operate at one of three clock frequencies ($L_{CLK}\{0, 1, 2\}$), with input data fed by pattern generators comprising a

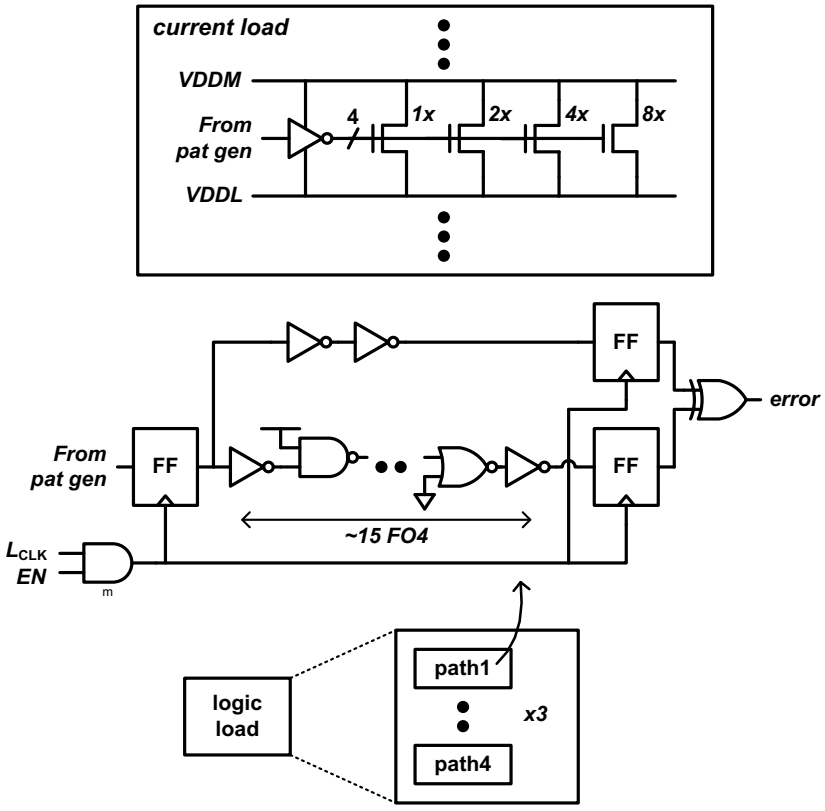


Figure 3.2: Detailed block diagrams of power-consuming *current load* and *logic load* blocks within a core

bank of linear feedback shift registers (LFSRs), programmable via scan chain. This enables configurations to (i) measure logic delay sensitivity to voltage and (ii) inject fixed or pseudo-random patterns of switching noise into the array at different layers. The *current load* block contains a bank of nMOS switches operating as programmable 4-bit binary-weighted current sources, resulting in a total of 45 *current load* settings per stack layer. This block models average current consumption of processor logic and the switches (operating in saturation) approximate the impedance across the rails of digital CMOS circuits. Separate pattern generators allow the *cores* to mimic larger magnitudes of current fluctuations and coarser-grained switching activity (e.g, clock

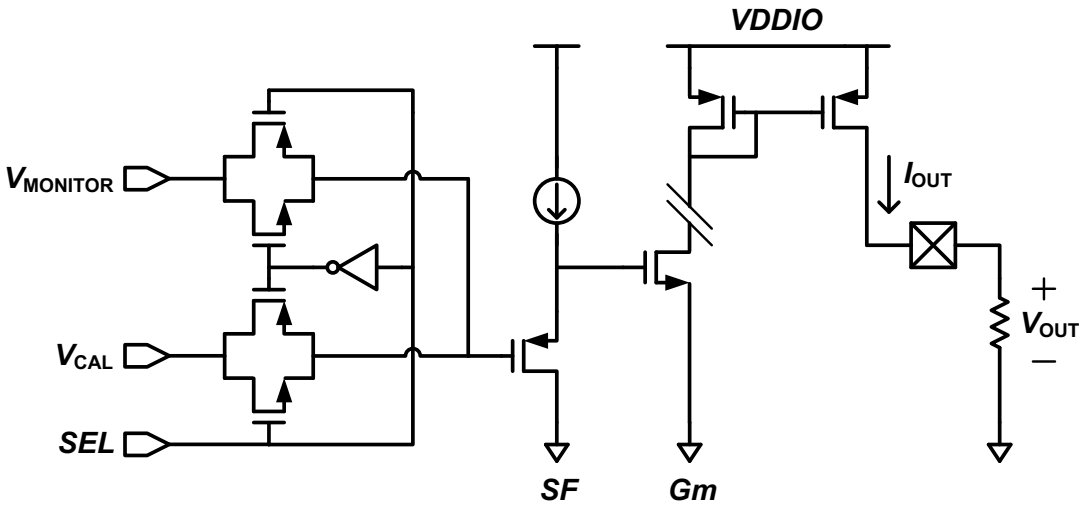


Figure 3.3: Block diagram of voltage monitor circuit for real-time voltage measurements

gating) at one of three clock frequencies ($I_{CLK}\{0, 1, 2\}$).

Voltage monitor circuit, shown in Figure 3.3, is used to capture the internal rail voltages in real time. The voltage monitor circuit consists of a source follower (SF) stage that senses the rail voltage followed by a transconductance (Gm) stage that converts the source follower output voltage into I_{OUT} [41]. The test-chip contains four separate voltage monitor circuits for each internal voltage rail (VDDH, VDDM, VDDL, VSSC).

The test chip was fabricated in MIT Lincoln Laboratory's (MITLL) 150nm FDSOI CMOS process. An annotated die photo of the chip is shown in Figure 3.4. Not only is voltage stacking easier with SOI process due to the absence of body effect, ultra-thin

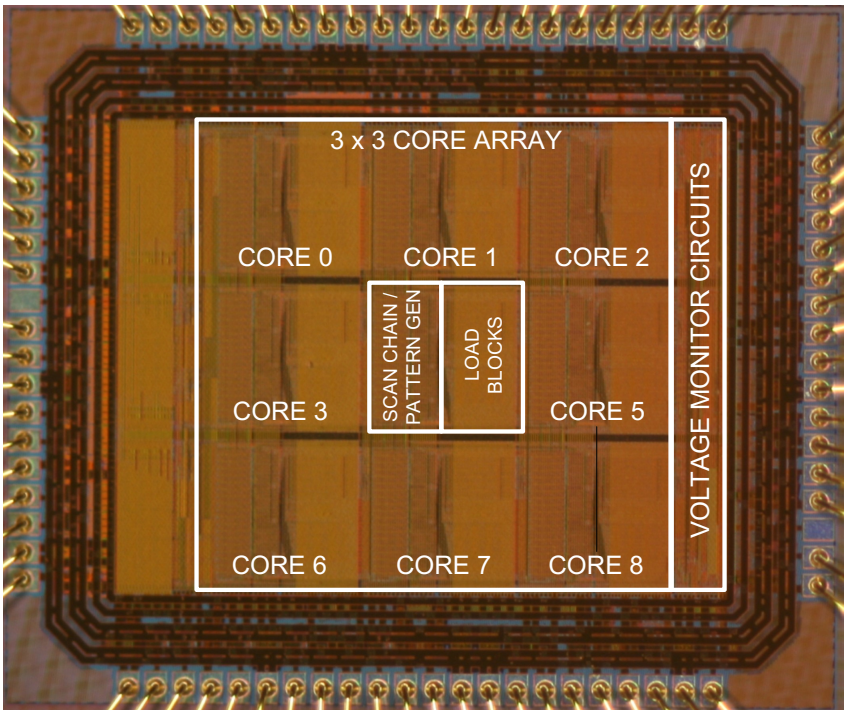


Figure 3.4: Annotated die photo of test chip prototype fabricated in MIT Lincoln Lab's 150nm FDSOI process.

body SOI technologies are one of the candidates for future aggressively scaled process nodes beyond 32nm. Hence, this test-chip encompasses key attributes to evaluate the potential of using voltage stacking in future designs. The MITLL FDSOI process was optimized for sub-threshold operation, with the nominal threshold voltage of the transistor at 0.4V. Hence, the test chip was designed to explore voltage stacking in the context of near-threshold multi-core computing at the intended core operating voltage of 0.5V. However, the findings presented in this chapter applies equally well to super-threshold and sub-threshold designs.

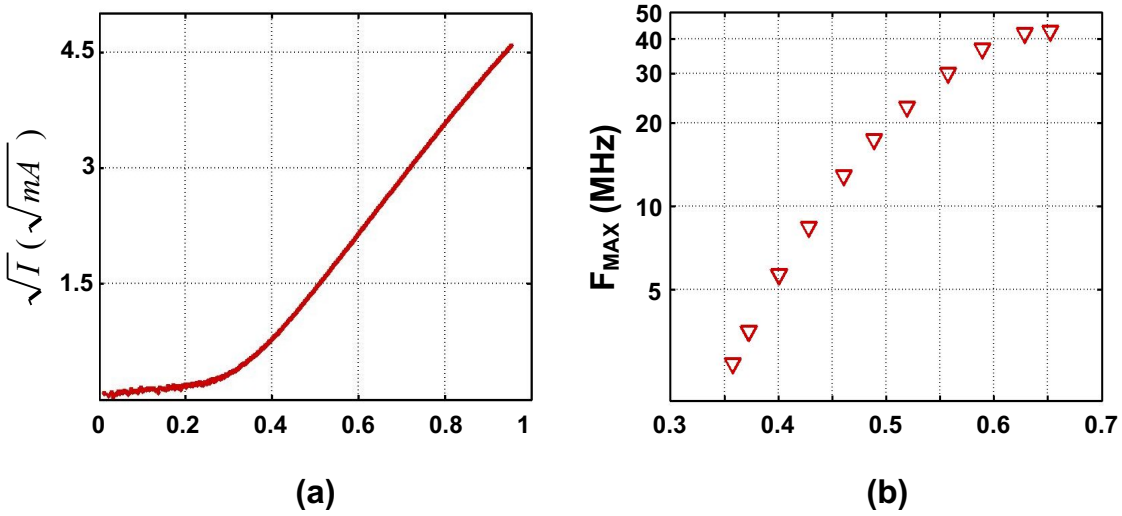


Figure 3.5: (a) *Current load* block's measured current and (b) *logic load* block's F_{MAX} vs voltage

3.2 Voltage Noise due to Static Activity Mismatch

Measured current versus voltage relationship for the 8x nMOS switch of the *current load* is presented in Figure 3.5(a). The maximum operation frequency (F_{MAX}) of a logic path in the *logic load* versus voltage across the stacks is shown in Figure 3.5(b). Exponential dependence of F_{MAX} on voltage (Figure 3.5(b)) clearly illustrate that the *cores* operate in the near-threshold regime at the intended *core* operating voltage of 0.5V.

Under static conditions, balanced inter-layer current consumption evenly distribute the stack voltage across the stack layers as dictated by KCL. Figure 3.6 plots the measured static array current (I_{ARRAY}) and the voltage levels that the internal stack voltages settle to, all versus identical inter-layer static *current load* settings, as the *current load* setting is increased from 0 to 45. For identical *current load* settings,

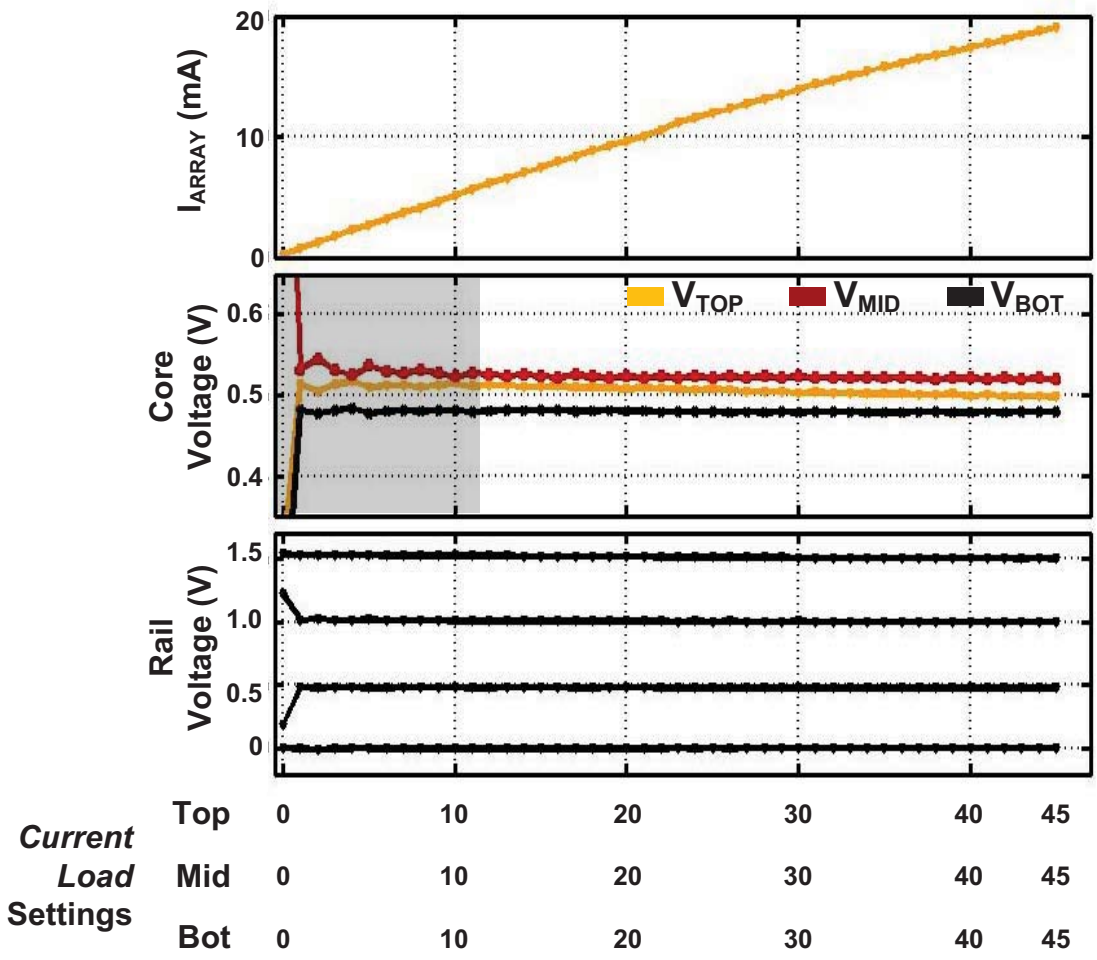


Figure 3.6: Measured static array current and voltage levels of voltage-stacked array for balanced *current load* setting

the array evenly splits VDDH (at 1.5V) into ∼0.5V levels. Inherent mismatches between the layers cause small disparities between core voltages. Furthermore, close examination of core voltages at low *current load* settings (shaded in gray) show that voltage deviations are larger at low *current load* setting; the largest difference seen when all blocks only leak (settings of 0). This is due to the characteristics of the internal rail impedance, which is further explained later on in the paper.

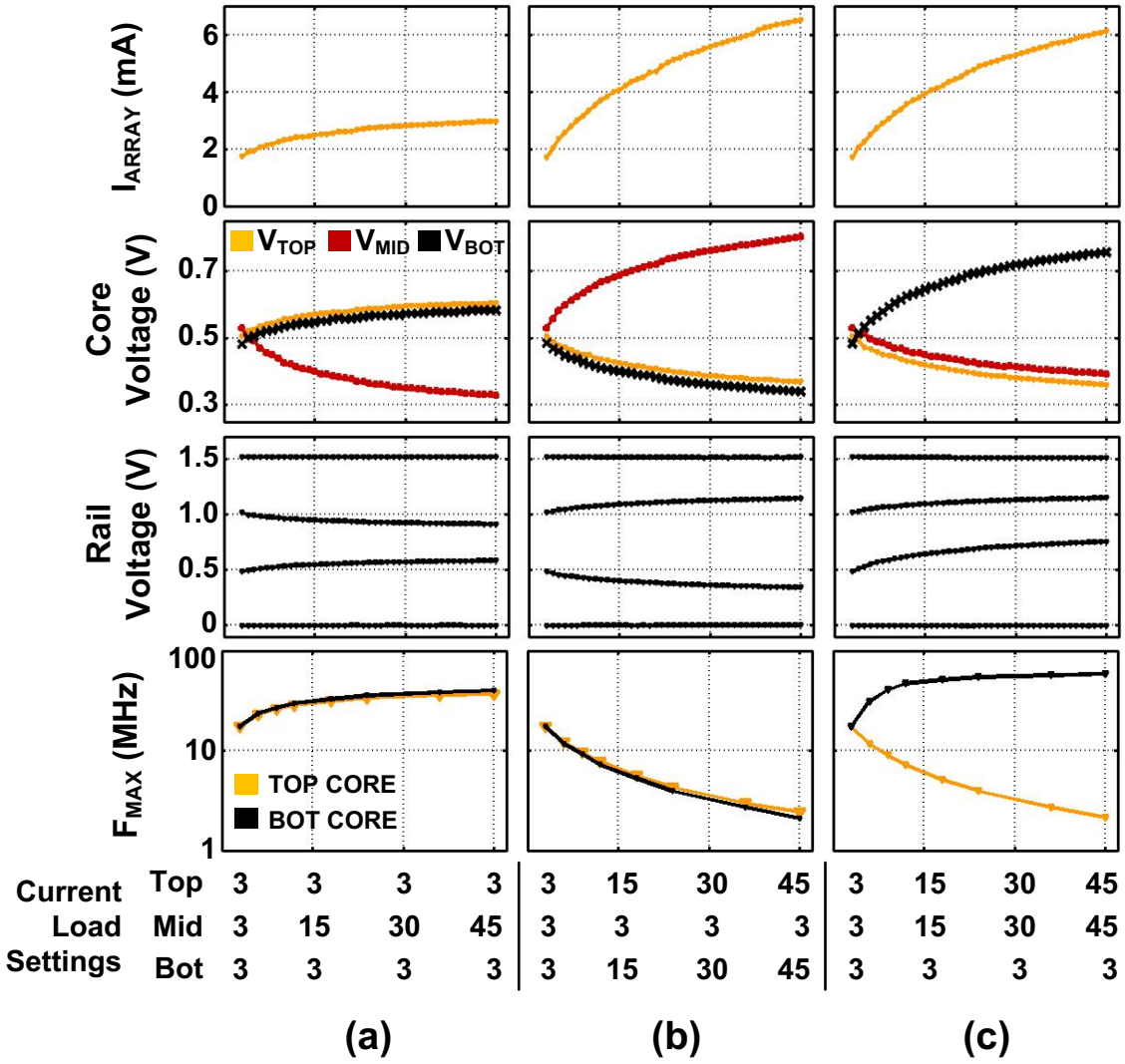


Figure 3.7: Measured static array current, voltage levels and F_{MAX} of voltage-stacked array for skewed *current load* settings with current increased only in (a) middle layer (b) top and bottom layers (c) top and middle layers

Figure 3.7 plots the measured static array current, the voltage levels and F_{MAX} of logic paths in the top and bottom layers, all versus various static skewed *current load* settings. A bug in the level-shifter circuit prevented measurement of F_{MAX} for logic paths in the middle layer. Figure 3.7 shows that skewed *current load* settings

between layers leads to uneven distribution of voltages. For example, as *current load* settings in the middle layer increase relative to the top and the bottom, held at a setting of 3 (Figure 3.7(a)), voltage across the middle layer decreases. The results show that to maintain balanced voltage distribution across the stack layer as shown in Figure 3.6, inter-layer activity difference must be absorbed using an active regulation scheme, such as push-pull regualtors [2], linear regulators [46] or switched-capacitor circuits [11]. However, using voltage regulators to absorb large inter-layer activity difference can be costly for sustained mismatches, as it incurs voltage regulator losses that degrade the system energy efficiency, .

Conversely, Figure 3.7 also demonstrates how processor cores within a layer could operate at different voltage/frequency settings from cores at different layers by intentionally skewing inter-layer power consumption. This could be achieved by enabling/disabling different number of cores per layer, essentially trading off core count for higher performance in a particular layer. Voltage stacking shows the potential to achieve workload heterogeneity while obviating voltage regulators.

To fully understand how internal voltages change with current mismatch, we investigate voltage deviation versus current mismatch relationship using a simple two-stack layer case illustrated in Figure 3.8, where external voltages VDD_{EXT} and GND_{EXT} are fixed and only V_{INT} change with activity mismatch. Starting from a balanced *current load* setting ($V_{TOP} = V_{BOT} = V_{BAL}$ and $I_{TOP} = I_{BOT} = I_{BAL}$), if *current load* setting of the top layer increase as shown in Figure 3.8(a), internal voltage V_{INT}

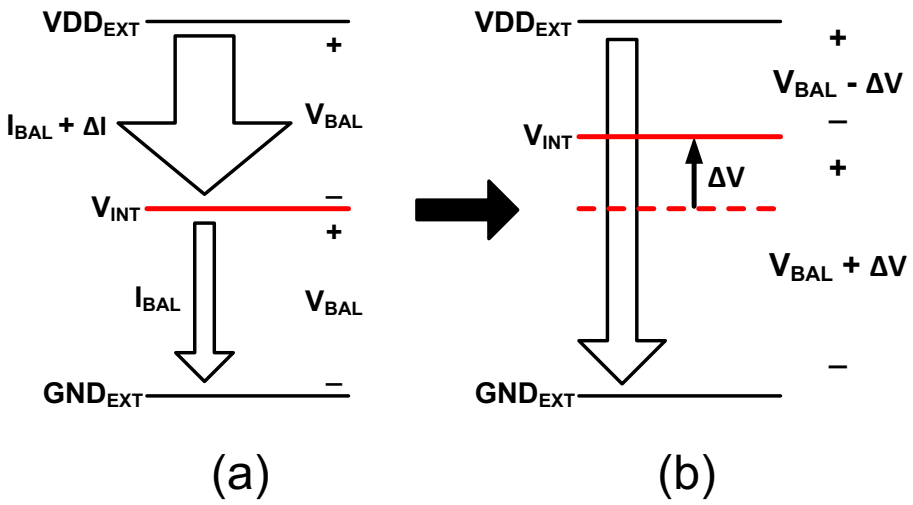


Figure 3.8: Diagram modeling voltage deviation for a simple two stack layer system

increases by ΔV to balance the current flow through the stack. Assuming power consuming blocks in each layer are nMOS switches like in the test-chip prototype, the current versus voltage relationship can be modeled using generalized alpha-power law

$$I = K \cdot (V)^{\alpha} \tag{3.1}$$

where K represents aggregate width of the transistor and we have ignored V_{TH} for simplicity. Increase in workload activity in the top layer as shown in Figure 3.8(a) can be modeled as

$$(K + \Delta K) \cdot (V_{BAL})^{\alpha} \tag{3.2}$$

where ΔK is the increase in aggregate nMOS device width in the top layer, which represents increase in activity. ΔV due to activity mismatch ΔK can then be calcu-

lated by solving the KCL equation

$$(K + \Delta K) \cdot (V_{BAL} - \Delta V)^{\alpha} = (K) \cdot (V_{BAL} + \Delta V)^{\alpha} \tag{3.3}$$

and substituting ΔI expression for V_{BAL}, which results in

$$\Delta V = \frac{\sqrt[\alpha]{K + \Delta K} - \sqrt[\alpha]{K}}{\sqrt[\alpha]{K + \Delta K} + \sqrt[\alpha]{K}} \cdot \sqrt[\alpha]{\frac{\Delta I}{\Delta K}} \tag{3.4}$$

Equations (3.1)−(3.4) can easily be expanded to other voltage stacked cases with more layers. Equation (3.4) illustrates how internal voltage deviations in voltage stacked systems relate to mismatch in current. For nMOS switches of with quadratic current versus voltage relationship ($\alpha \approx 2$) (Figure 3.5(a)), the shape of VDDM and VDDL curves result in square root shape as seen in Figure 3.7.

The equations also apply to synchronous CMOS logic blocks ($I \propto CVF$), where K in equation (3.1) corresponds to aggregate switched capacitance of the logic block, rather than transistor width. Therefore, for synchronous digital logic Equation 3.3 becomes:

$$I_{\text{STACK}} = \alpha \cdot C \cdot f \cdot (V_{\text{BAL}} + \Delta V) = (\alpha + \Delta\alpha) \cdot C \cdot f \cdot (V_{\text{BAL}} - \Delta V) \tag{3.5}$$

where $\alpha \cdot C$ is the total switched capacitance, α is the activity mismatch, f is the clock frequency, ΔV is the magnitude of voltage deviations in the stack layers, and I_{STACK} is the current flow through the stack layers.

Equations (3.3)-(3.5) reveal intuitive yet important insights into the determining factors of noise in voltage stacked systems:

- **Activity mismatch ($\Delta\alpha$):** Layer voltages deviate from V_{BAL} to compensate for the activity mismatch between stack layers and maintain identical current through the stack. Equation (3.5) shows that larger activity mismatch $\Delta\alpha$ results in larger voltage deviations. For nMOS *current loads* implemented in the test chip, mismatch in transistor width (ΔK) corresponds to the activity mismatch in digital systems.

- **Current dependence on voltage:** In addition to activity mismatch $\Delta\alpha$, inter-layer voltage noise also depends on the magnitude of the nominal current flowing through the stack (I_{STACK}), set by α (or K for the test chip). Hence, smaller values of ΔV can neutralize the same $\Delta\alpha$ (or ΔK) when overall activity and power is higher. Moreover, if clock frequency, *f*, of each layer can vary proportionally with voltage, dynamic current has a quadratic, rather than linear, dependence on voltage similar to nMOS loads, which further reduces voltage noise. In other words, the overall power consumption and the choice of clocking strategy can both impact voltage noise on the stack layers.

3.3 Characteristics of Supply Rail Impedance in Voltage Stacked Systems

Voltage noise dependence on the current versus voltage relationship described in the previous section reveals an important underlying attribute of voltage stacking. Unlike traditional power delivery schemes where circuitry has low-impedance connection to fixed off-chip supply rails as shown in Figure 3.9 (a), circuitry in voltage

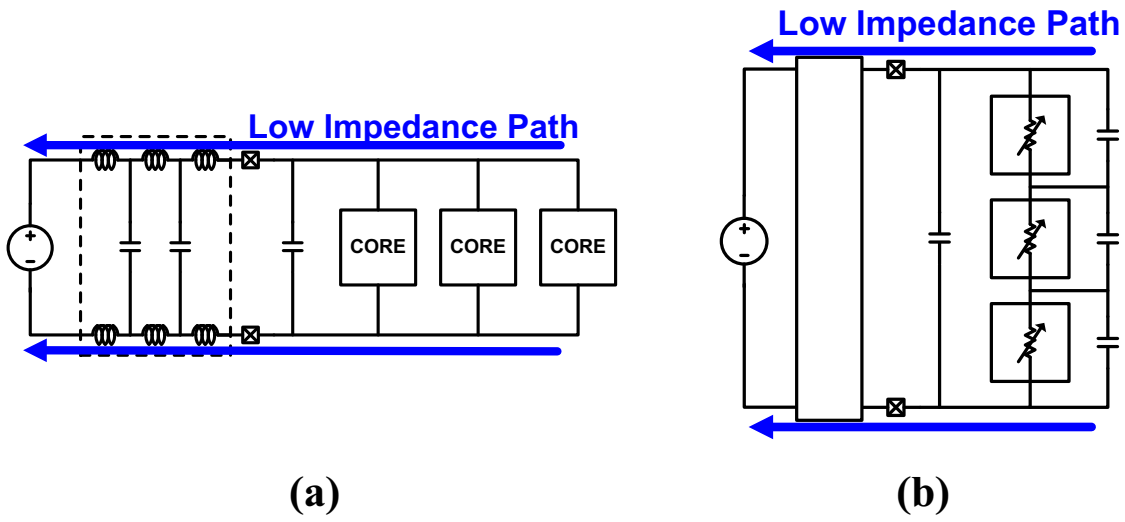


Figure 3.9: Comparison of power delivery impedance between (a) conventional power delivery system and (b) voltage stacked system

stacking suffer from relatively high impedance paths for one or both rails, depending on the layer, as illustrated in Figure 3.9 (b), because all cores connect to the external low impedance rails through the cores in other layers. For instance, for the core in the bottom layer, only the ground rail is connected directly to a low impedance path, while the VDD rail is connected to the ground rail of the core in the middle layer. As such, the impedance of the power supply rails of the core in the bottom layer depends on the power consumption of the core in the middle. Hence, the impedance of the internal supply rails of a stack layer depends on the power consumption of the other stack layers, and thus, the total current flowing through the stack determines the supply rail impedance of the stack layers. Therefore, the aggregate power consumption of the cores across the stack, represented by the overall current flow, has direct impact on voltage noise. At higher overall power (higher current flow through the stack), the supply rail impedance of the stack layers is lower. Therefore, any

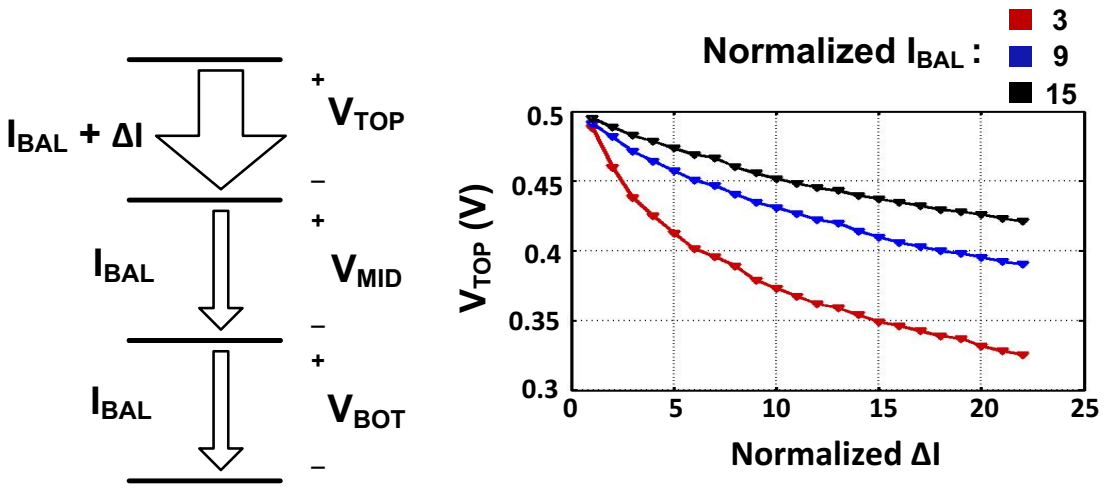


Figure 3.10: Measured V_{TOP} values for various initial power consumption conditions for inter-layer activity skew (*current load* setting increased only in top layer)

inter-layer activity mismatch between the layers, which would cause change in the overall current flow, results in smaller voltage deviations, as the incremental change in current occurs through a lower impedance path.

To demonstrate this property, Figure 3.10 plots the voltages that V_{TOP} settle to versus the normalized differences in inter-layer *current load* settings when the *current load* settings are skewed by increasing only the setting of the top layer, for three different values of balanced current I_{BAL}. Higher initial I_{BAL} values (higher chip power consumption), result in smaller voltage deviation for the same inter-layer activity mismatch, due to lower impedance of the internal supply rails. This can also be seen from equation (3.4) and (3.5), where higher K (or higher α) value translates to higher current consumption (for the same voltage conditions), leading to a smaller ΔV for the same mismatch ΔI (ΔK). This suggests that voltage stacked systems become more immune to inter-layer activity mismatch as overall power consumption

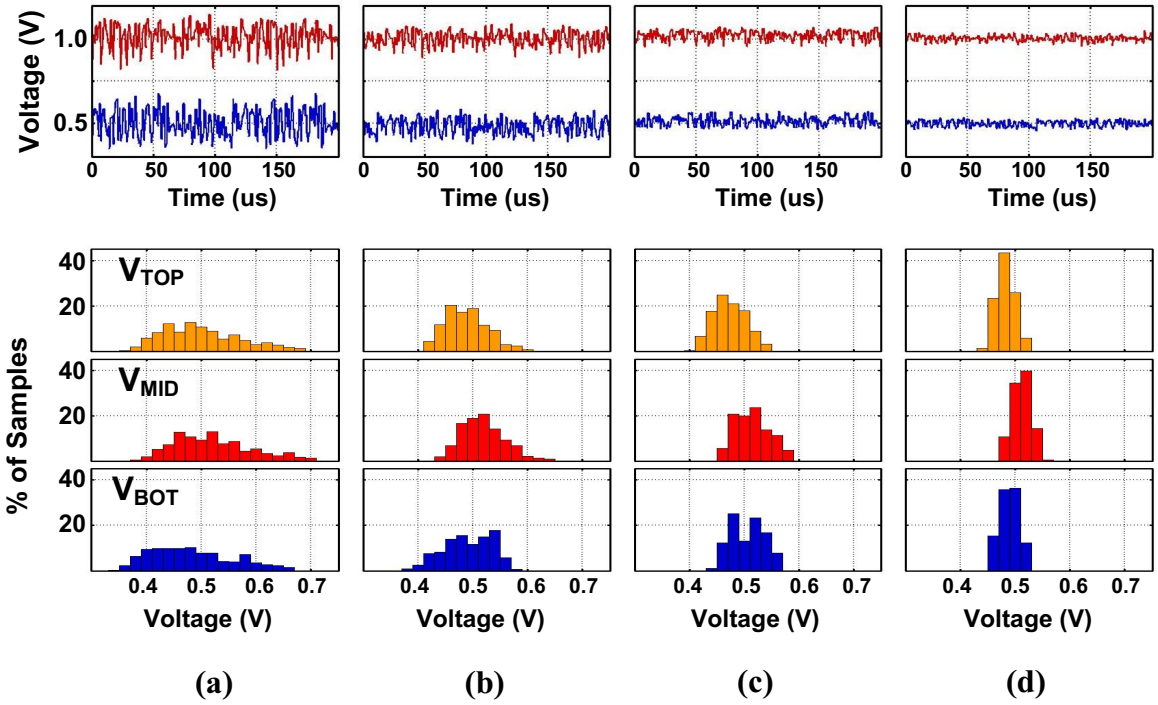


Figure 3.11: Measured VDDM and VDDL voltage traces (top) and histogram of V_{TOP}, V_{MID} and V_{BOT} when pseudo-random fixed-magnitude noise is injected in all three stack layers for balanced *current load* setting of (a) 3, (b) 8, (c) 15, and (d) 31

is increased.

This dependence is also illustrated in Figure 3.11 by measuring the internal voltage variation in the test chip when pseudo-random noise is injected into the array for different balanced *current load* settings. Noise is injected by toggling the remaining *current load* bits. Real-time internal voltage trace of $VDDM$ and $VDDL$, along with histogram plots of V_{TOP}, V_{MID} and V_{BOT} are shown. Pseudo-random noise of fixed magnitude was injected into all three stack layers for four different balanced *current load* settings. Figure 3.11 clearly shows that as power consumption is increased, voltage stack becomes more robust to noise, demonstrating that the voltage variation

depends on activity mismatch and overall power, and not the absolute value of activity mismatch. This attribute suggests that as we continue to integrate more cores on a single chip, the power fluctuation of a single core due to architectural events such as cache miss will have lesser impact on internal voltage for the same power in future voltage stacked many-core systems. Furthermore, for many-core systems, the worst case voltage noise scenario in which all cores in a stack layer experience power fluctuations at once, can be easily avoided by workload scheduling via software. This suggests that for many-core voltage stacked systems, coordinated workload scheduling could provide efficient voltage regulation during high-power states while on-die regulators may only be relied on during low-power states.

The experimental results presented in this chapter explored the basic characteristics of voltage noise due to inter-layer activity mismatch in voltage stacking, and demonstrated that the internal supply rail impedance in voltage stacked systems depends on aggregate power consumption of the stack layers unlike conventional power delivery schemes. Thus voltage noise depends on the activity mismatch to overall aggregate activity ratio rather than the absolute magnitude of the mismatch. Based on this characteristics, we can deduce that voltage noise can be mitigated by either i) reducing the inter-layer activity (power) mismatch or ii) by operating the system at a higher overall power. These results suggest that voltage stacking will scale well with future high-throughput many-core systems, where power fluctuations in a single core has less impact on voltage noise for the same overall power consumption, and there is more opportunity for software-level workload balancing between the stack layers.

Chapter 4

Mitigating Inter-layer Voltage Noise using Adaptive Clocking and Fully-Integrated Voltage Regulator

As shown in the previous chapter, voltage noise in voltage stacked systems depends on how well the workloads are balanced across the stack layers, suggesting that software-level workload balancing can be an effective deterrent against voltage noise, especially for high-throughput systems. However, the timescales of such software techniques to mitigate inter-layer activity mismatch are usually too long. It is also infeasible to rely wholly on software techniques to provide operating voltage guarantees necessary to avoid noise-related failures such as SRAM instability or timing violations. Moreover, software techniques cannot easily address situations where entire layers must be powered down. Hence, we explore using efficient voltage regulators to neutralize activity mismatch and mitigate internal voltage noise in voltage

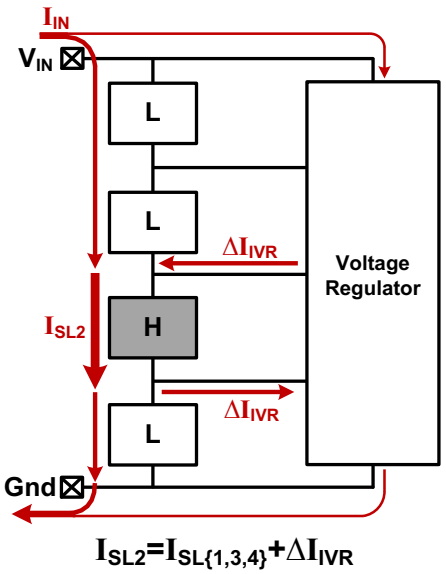


Figure 4.1: Block diagram illustration of voltage regulation in voltage stacking. The integrated voltage regulator (IVR) provides additional power to support the higher activity in the core in stack layer 2.

stacked systems. Fig. 4.1 illustrates how an on-chip integrated voltage regulator can neutralize inter-layer voltage noise in voltage stacking. In this scenario, the core stack layer 2 (SL2) executes a higher activity program **H** compared to the cores in the other three layers (all running a lower activity program **L**). To mitigate the voltage droop in SL2, voltage regulator provides the additional current required to support the higher activity workload in the higher-activity core. Notice the IVR only provides the differential current (ΔI_{IVR}). Therefore, voltage regulator losses only apply to a small fraction of the overall power delivered to the stack, and the overall on-chip power delivery efficiency is not limited by the IVR efficiency.

This chapter explores efficient voltage noise mitigation via an integrated voltage regulator (IVR) with processor cores operating in adaptive frequency clocking mode, demonstrated using a 16-core test chip prototype that integrates industry-

grade microprocessor cores with a symmetric-ladder switched-capacitor (SC) DC-DC converter. Building upon prior works on voltage regulators for voltage stacking applications presented in Section 4.1, Section 4.2 presents the implementation details of a test chip prototype fabricated in TSMC's 40G process. Next, Section 4.3 presents measurement results that demonstrate robust noise mitigation benefits of adaptive frequency clocking and the IVR, and the impact on system performance and energy-efficiency. Section 4.4 then takes a step back to understand the improvements in system-wide power delivery efficiency made possible by voltage stacking. Finally, Section 4.5 explores how intelligent workload allocation at the software-level can augment the hardware techniques to further improve the system efficiency by exploring the impact of wider workload diversity and the impact of workload allocation strategies in a multi-core voltage-stacked system.

4.1 Prior works on Noise Mitigation in Voltage Stacking

There has been several examples of prior work that have proposed different voltage regulator topologies to regulate the stacked layers in a voltage-stacked system. Push-pull linear regulators have been proposed to handle the inter-layer activity mismatch [47, 46]. Although linear regulators have small area overhead and are easy to integrate, they suffer from inherently low conversion efficiency, especially for high step down ratios. For $n-$way voltage-stacked system, the power delivery efficiency is less than $1/n$ in the worst case scenario. Off-chip inductor based switching regulators have

been shown for voltage stacking applications [57, 28], but the inherent difficulty of integrating high-quality inductors on-chip hinders full integration of inductor-based regulators.

On the other hand, high quality capacitors are much easier to integrate, making switched-capacitor (SC) converters attractive. For example, [11] demonstrated a 2-to-1 SC converter that regulates the intermediate voltage of two stacked layers. Multiple 2-to-1 SC converters have been used to support more than two stacked layers [73, 74] at the expense of higher complexity and overhead. Building on prior works, this chapter proposes using a symmetric ladder topology to implement a fully-integrated 4-to-1 SC converter that mitigates voltage noise on four stack layers simultaneously.

The remainder of the chapter presents the design of a 16-core voltage-stacked test chip that demonstrates efficient voltage noise mitigation using not only a switched-capacitor integrated voltage regulator, but by exploring adaptive frequency clocking (AFClk) scheme for the cores in the voltage stack, thereby providing a holistic approach to mitigating voltage noise in voltage stacking. We evaluate the noise mitigation and performance advantages the IVR and AFClk provide, and demonstrate the advantages of voltage stacking for efficient power delivery in multicore systems.

4.2 Overview of 16-core Voltage Stacked System Architecture

Figure 4.2 presents an overview of the test chip implemented in TSMC's 40G process [34]. It comprises 16 processor cores organized into a 4x4 stacked array. The core

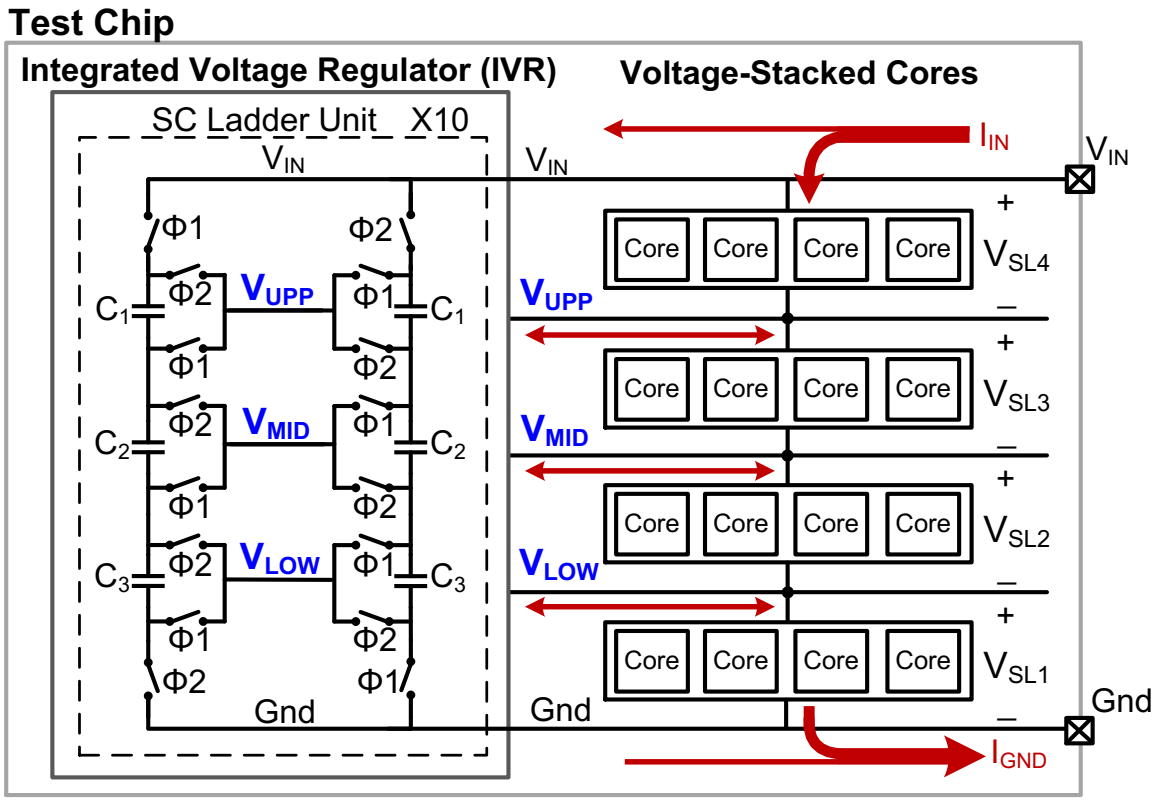


Figure 4.2: Block diagram of the voltage-stacked system showing the 16 4-way stacked cores and symmetric ladder switched-capacitor IVR.

array operates off a single 3.6V V_{IN}, with cores in each stack layer (SL{1:4}) operating nominally at layer voltages ($V_{SL\{1:4\}}$) of 0.9V, which is the nominal operating voltage for this process. Each layer relies on triple wells to isolate nMOS transistors in each layer from body-bias effects. As described in the previous chapter, when power consumption in all layers perfectly matches, voltage stacking evenly distributes internal voltage rails to 0.9V per layer. However, as only V_{IN} is connected to a fixed external supply rail, the internal rails V_{UPP}, V_{MID}, V_{LOW} fluctuate when there is inter-layer activity mismatch between the cores occupying different layers, resulting in voltage noise. The test chip implements adaptive frequency clocking (AFClk) for the cores

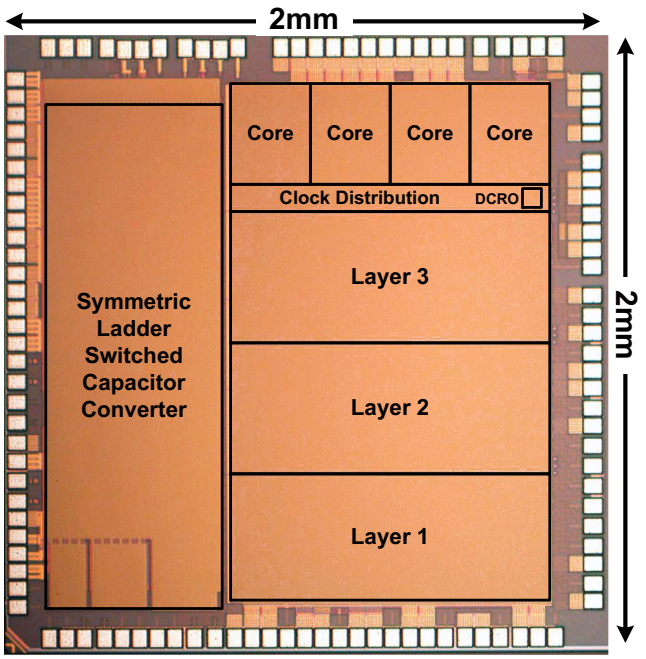


Figure 4.3: Annotated die photo of the test chip implemented in TSMC 40G process.

in each layer and integrates an IVR on the same die to reduce inter-layer voltage noise. The test chip also integrates additional circuitry for testing and debug purposes: layer-shifting circuits for signaling between different stack layers; scan chain to configure the digital blocks; and voltage monitoring circuitry to probe internal voltage rails and measure real-time voltage noise. An annotated die photo of the test chip is shown in Figure 4.3.

4.2.1 Integrated Voltage Regulator

The test chip integrates a 4-to-1 switched-capacitor (SC) IVR with the core array to neutralize voltage noise due to activity mismatch between the stack layers, as shown in Figure 4.2 [34, 61]. The IVR implements a symmetric-ladder topology consisting of

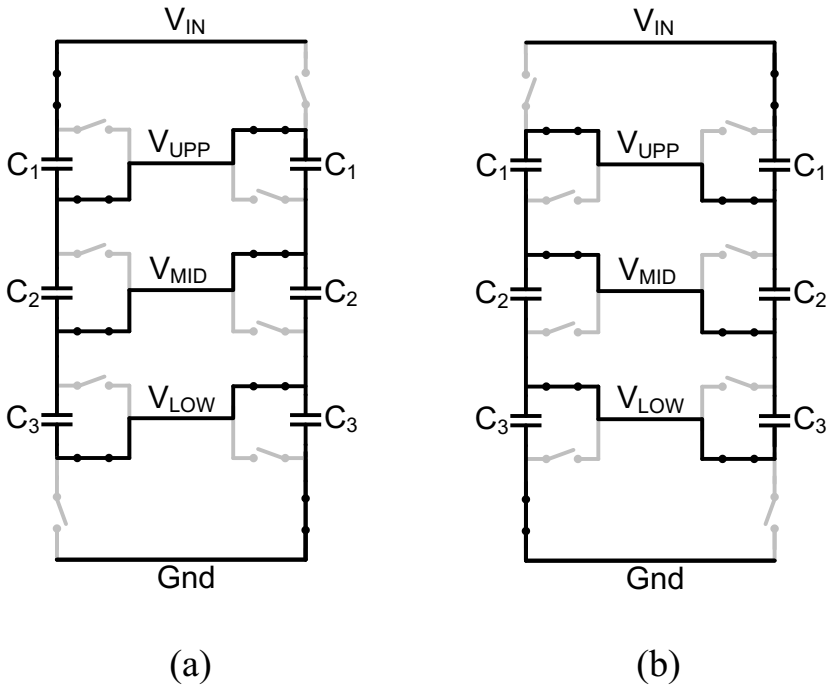


Figure 4.4: Illustration of two-phase operation of the IVR: (a) Phase 1 (b) Phase 2.

10 SC ladder units, each controlled by one of 10 phase-interleaved switching signals to reduce voltage ripple. The symmetric-ladder topology is a natural choice for voltage stacking, because in a sense, it too employs voltage stacking. By connecting to the internal rails V_{UPP}, V_{MID}, and V_{LOW}, the 4-to-1 symmetric-ladder converter becomes a multi-output regulator that pushes and pulls current to and from the stacked core array to smooth out imbalances and reduce voltage noise in the four stacked layers simultaneously. Because the IVR only neutralizes activity *mismatch* between the layers, the maximum power the IVR would ever need to provide is the power consumed by four cores in a single layer. It is important to note this is one quarter of the power a conventional single-layer, 16-core system would require from an IVR.

Powered from the external 3.6V V_{IN}, the SC ladder unit operates with respect

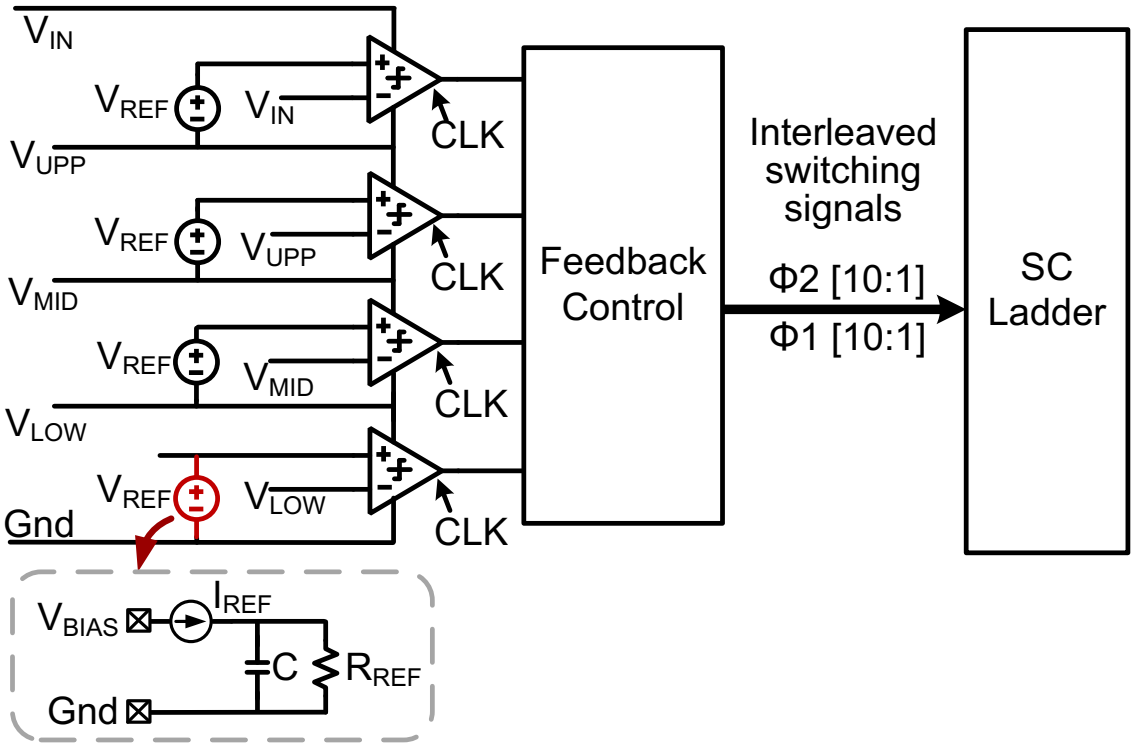


Figure 4.5: Implementation of IVR lower-bound feedback control.

to two non-overlapping clocks Φ1 and Φ2, as shown in Figure 4.4. In phase Φ1, the left capacitor ladder is connected to V_{IN} while the right capacitor ladder is connected to ground (Figure 4.4(a)), while in phase Φ2, the right ladder is connected to V_{IN} and the left ladder is connected to ground in a symmetric fashion. The operation of the IVR is similar to other SC based converters: the flying capacitors are charged in one phase and discharged in the other in symmetric fashion while current flows from the input to the output nodes through the capacitors and the power switches. Due to the ladder structure, The power switches and the flying capacitors in the main SC ladder can be implemented with thin-oxide NMOS or PMOS transistors, as the voltage across them are always around the nominal operating voltage.

Rather than regulating the internal rails to a specific voltage, the IVR implements a single-bound feedback control to keep each layer voltage above a prescribed reference voltage, as illustrated in Figure 4.5. Each layer implements a 2.5GHz digital clocked comparator circuit that compares the layer voltages $V_{SL\{1:4\}}$ to a reference voltage V_{REF} generated on-chip for each layer. If any of the layer voltages fall below V_{REF}, the associated comparator generates a pulse signal that is processed by the feedback control logic to produce the clock signals Φ1 and Φ2 for each of the 10 SC ladder units, which switch at a maximum frequency of 250MHz. As a result, the IVR operates to keep the lower bound of each layer voltage at V_{REF} for all stack layers. The reference voltage V_{REF} is created on-chip by reference current flowing through a resistor, as shown in Figure 4.5, so that

$$V_{\mathrm{REF}} = I_{\mathrm{REF}} \cdot R_{\mathrm{REF}} \tag{4.1}$$

I_{REF} is tuned using an off-chip voltage source V_{BIAS}, which creates a stable reference voltage. The parallel coupling capacitor C couples high-frequency noise of the internal rails to the input of the comparators to overcome the slow slew rate limits of the resistor.

One thing to point out, however, is that SC converters cannot regulate the stack layers to an exactly even distribution of V_{IN}. This is because the SC converter relies on a certain amount of voltage difference between the output and the flying capacitors to deliver charge to the output. N-to-1 SC converters typically regulate the output voltage to (V_{IN}/N - ΔV), and the converter efficiency and it's ability to deliver charge diminishes as ΔV approaches zero [32]. To overcome this limitation

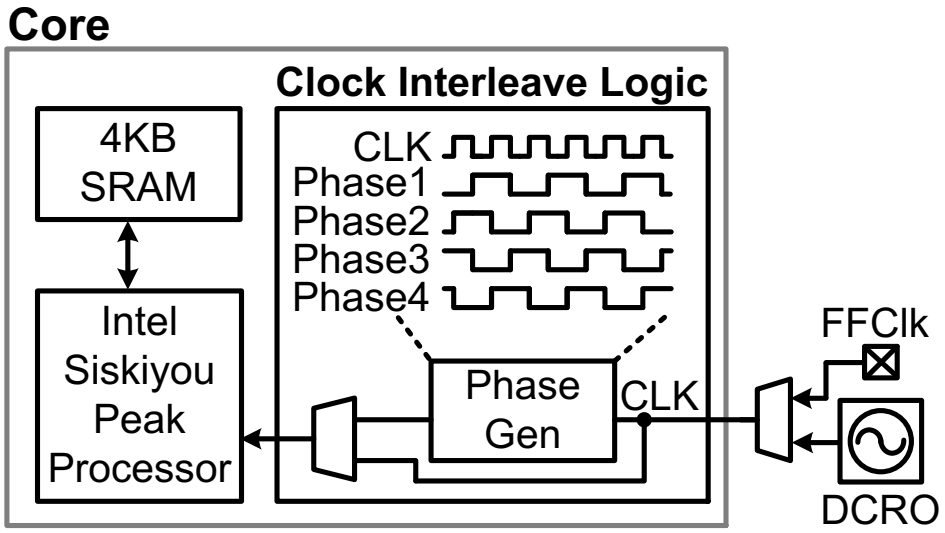


Figure 4.6: Block diagram of the core implementing Intel's Siskiyou Peak processor with 4KB SRAM and two modes of clocking.

and augment the SC converter to improve system efficiency, this chapter also explores per-layer adaptive frequency clocking (AFClk), where clock frequency of cores in each stack layer tracks the fluctuations in layer voltage. Adaptive frequency clocking is particularly a good fit with voltage-stacked systems. By having the per-layer clock frequency track voltage fluctuations, current consumption has a stronger dependence on voltage, which further alleviates voltage noise.

4.2.2 Core Design

Figure 4.6 presents a block diagram of the *core* design. Each core contains an Intel Siskiyou Peak processor with 4KB SRAM for both instruction and data. Our version of the Siskiyou Peak processor is a 5-stage, single-issue, integer pipeline that implements a subset of the IA instruction set architecture and system software model. Each core can be independently programmed for multi-program operation. To inves-

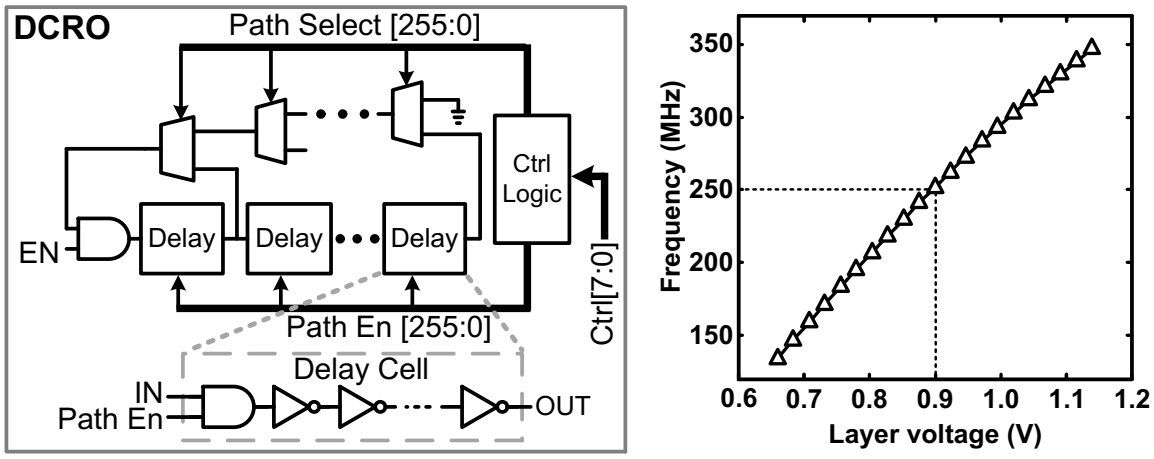


Figure 4.7: Block diagram of the per-layer Digitally Configurable Ring Oscillator (DCRO) (Left). Measured DCRO voltage vs. frequency (Right).

tigate the interaction between voltage noise and clocking schemes in voltage-stacked systems, the cores can operate in one of two clocking modes: (i) fully synchronous global fixed-frequency clocking(FFClk) or (ii) per-layer adaptive frequency clocking (AFClk). For fixed-frequency clocking, an external clock source drives all 16 cores. Per-layer adaptive frequency clocking utilizes a digitally configurable ring-oscillator (DCRO) in each layer to generate the per-layer clock, whose frequency tracks the fluctuations in the stack layer voltage.

4.2.3 Adaptive Frequency Clocking

Prior works have demonstrated numerous adaptive frequency clocking schemes that utilize critical path tracking circuitry to change the clock frequency with voltage fluctuations [13, 18, 35, 76, 60]. The DCRO shown in Figure 4.7, comprises a programmable array of delay cells configured via an 8-bit control code with 125ps per bit delay resolution. DCRO clock period acts as a proxy for the critical path delay

of the cores, and similar designs have been shown to deliver good tracking accuracy [66, 8, 75, 23]. The DCRO is configured such that the cores operate at the maximum frequency without timing violations across the entire operating voltage range of 0.65 to 1.15V. The frequency versus voltage relationship of the DCRO at the setting corresponding to the maximum core frequency is presented in Figure 4.7. The DCRO runs at a nominal frequency of 250MHz and consumes 2.31mW at 0.9V.

4.2.4 Clock-phase Interleaving

In addition to adaptive frequency clocking, each layer implements clock-phase interleaving (ClkInt) that provides each of the four cores in each layer with 90° out-of-phase clock signals to smooth out high-frequency within-cycle voltage noise/ripple.

4.2.5 Layer-shifter circuit

While the cores do not communicate with each other in this test chip, inter-layer communication is still required to transmit clock signals to and from the cores, and to send scan signals across the stack layer boundaries. The test chip implements a capacitively-coupled, initializable, layer-shifter circuit to communicate between different stack layers. Figure 4.8 presents the implementation of a layer shifter circuit that shifts signals from SL2 to SL3 as an example. To enable domain crossing between any two stack layers without breakdown concerns, the coupling capacitors are implemented using metal-oxide-metal (MoM) capacitors. The sizing of the coupling capacitors provides a trade-off between power, reliability, area, and speed. Larger capacitors provide stronger coupling between layers, but also lead to larger area and

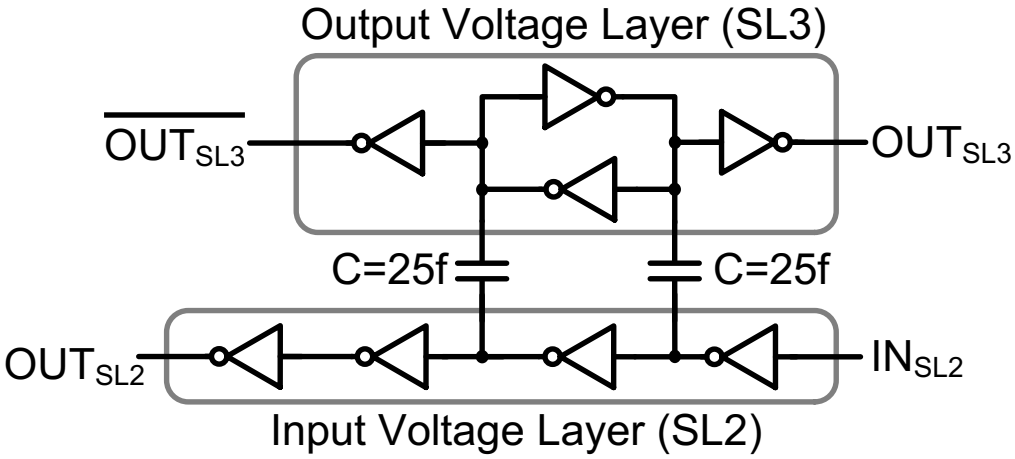


Figure 4.8: Implementation of initializable capacitively-coupled layer shifter circuit. Schematic illustrates layer shifter circuit conmmunicating data from SL2 to SL3. frequency (Right).

Table 4.1: Instruction power characterization

Instruction	Description	Power per core(@250MHz/0.9V)
IMUL	Signed multiply	14.27 mW
ADC	Add with carry	10.02 mW
XCHG	Exchange data	5.65 mW

power penalties. The insertion delay of the layer-shifter is less than 100ps. Similar circuitry was demonstrated in [46] using MOS capacitors.

4.3 Measurement Results: Voltage Noise Mitigation

This section provides the measurement results from the test chip fabricated in 40nm CMOS. The section begins with a brief discussion of the various workloads that were characterized to run on the cores and used for all of the subsequent measure-

ments. The remaining subsections analyze voltage noise in the 4x4 voltage-stacked test chip and evaluate the noise mitigation offered by the IVR and adaptive clocking. Section 4.3.2 characterizes the inherent inter-layer voltage noise for balanced and unbalanced workload scenarios assuming a single clock domain for all 16 cores and the IVR disabled. Section 4.3.3 then evaluates the noise mitigation provided by turning on the IVR, which reduces worst-case voltage droop. To further improve efficiency, Section 4.3.4 explores the benefits of adaptive frequency clocking.

4.3.1 Stressmark Generation

As discussed in Chapter 3, noise in voltage-stacked systems is a direct result of activity mismatch between the stack layers. The magnitude of voltage noise ultimately depends on: (i) workload behavior, (ii) the ISA and microarchitecture of the cores, and (iii) the underlying process technology that determine the leakage and dynamic power consumption of the cores in the system. Therefore, to profile the worst-case voltage noise of the test chip, the first step is to generate the power usage profile of the instructions in the ISA to find the maximum and minimum power instructions, and determine the worst-case nominal power difference that can be generated by the cores. This is done by generating a power-profiling micro-benchmark for each instruction that runs an endless loop of a 4KB code block that runs each instruction in the ISA continuously with no dependencies between instructions. Table 4.1 presents the measured dynamic power consumption per core for the maximum (IMUL), minimum (XCHG), and medium (ADC) power instructions of the ISA. It should be noted that the power usage changes depending on the operand type for each instruction. There-

Table 4.2: Micro-benchmarks used for noise profiling

Micro-benchmark	Description
MAX-CONT	IMUL instructions with random inputs in endless loop
MIN-CONT	XCHG instructions with random inputs in endless loop
MAX-MIN-n	n IMUL instructions and n XCHG instructions in endless loop
MID-MIN-n	n ADC instructions and n XCHG instructions in endless loop

fore, different operand types were also profiled to find the maximum and minimum power instructions.

Using this instruction power profile, we generate several micro-benchmarks to stress the voltage stack. The micro-benchmarks used for instruction profiling are utilized to generate static activity differences between the cores. We also generate several simple micro-benchmarks that generate oscillatory activity behavior by periodically switching between two different instructions. The micro-benchmarks used for testing are summarized in Table 4.2.

4.3.2 Voltage Noise in Fixed-Frequency Clocking Operation

To demonstrate the inherent properties of voltage stacking in the context of high-throughput multicore systems, we first characterize voltage noise with all 16 cores active, running synchronously off of a fixed-frequency clock (FFClk), and with the IVR and clock-phase interleaving (ClkInt) turned off. Figure 4.9 presents snapshots of measured V_{SL2} for various workload scenarios as well as per-layer box plots (which delineates a distribution's minimum, first quartile, median, third quartile, and maximum values) of all four layer voltages. The box plots show voltage distribution measured over a 1ms execution window. To ensure all cores operate correctly with no timing

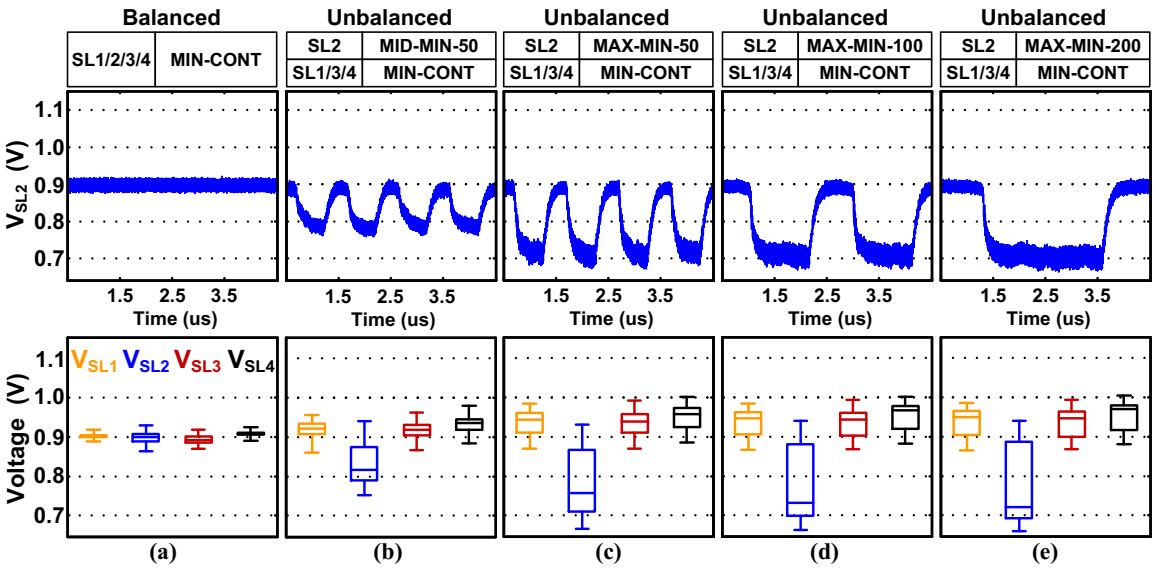


Figure 4.9: Measured transient waveforms of V_{SL2} and boxplots of V_{SL1}-V_{SL4} for balanced and unbalanced workload scenarios. (a) Balanced: All cores running MIN-CONT microbenchmark (b) Unbalanced: Cores in SL2 running MID-MIN-50 (c) Unbalanced: Cores in SL2 running MAX-MIN-50. (d) Unbalanced: Cores in SL2 running MAX-MIN-100. (e) Unbalanced: Cores in SL2 running MAX-MIN-200.

errors, the worst-case voltage noise, the corresponding minimum stack layer voltage V_{MIN}, and maximum operating frequency F_{MAX} were characterized for all cores. For all test results in this section, cores were run at the worst-case F_{MAX}=130MHz that guaranteed correct functionality under worst-case voltage noise.

Figure 4.9 compares the resulting voltage noise between a balanced workload scenario and multiple unbalanced scenarios. In the balanced scenario (Figure 4.9(a)), all cores ran the *MIN-CONT* micro-benchmark in lockstep. The V_{SL2} waveform and the boxplots show that all layers evenly subdivide to 0.9V, as expected, since all layers have identical activity. Figure 4.9(b)-(e) presents results for four unbalanced workload scenarios with the four cores in SL2 running a higher power micro-benchmark while all other cores ran *MIN-CONT*, demonstrating significant voltage droop in V_{SL2}.

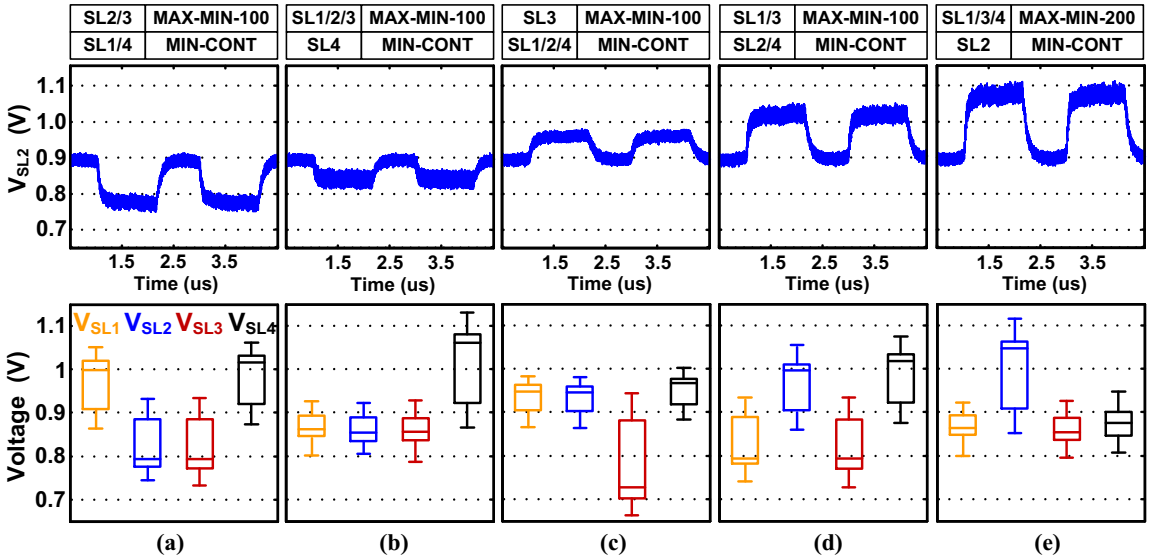


Figure 4.10: Measured transient waveforms of V_{SL2} and boxplots of V_{SL1}-V_{SL4} for various unbalanced workload scenarios: (a) Cores in SL2/SL3 running MAX-MIN-100 (b) Cores in SL1/SL2/SL3 running MAX-MIN-100 (c) Cores in SL3 running MAX-MIN-100 (d) Cores in SL1/SL3 running MAX-MIN-100 (e) Cores in SL1/SL3/SL4 running MAX-MIN-100.

Comparison of Figures 4.9(b) and (c) show worse voltage noise for larger inter-layer activity mismatch. Scenarios (c)-(e) present noise profiles for worst possible activity mismatch (*MAX* vs. MIN instructions), but with varying periodicity. Figure 4.9(c)-(e) show comparable noise magnitude for the three scenarios, while longer periods of activity mismatch results in longer periods of voltage droop akin to IR drop. The scenarios shown in Figures 4.9(c)-(e) all correspond to worst-case voltage noise when all 16 cores are active, as they generate: (i) the maximum possible activity mismatch between stack layers, and (ii) only a single layer has higher activity.

Since the voltage across a capacitor cannot change instantaneously, the rate of voltage transition due to inter-layer activity mismatch for any layer voltage depends on the time it takes to discharge the capacitance of the stack layers, which includes

the intrinsic capacitance of the digital logic, decoupling capacitance, and the non-switching flying capacitance of the IVR. The capacitance is discharged through the power grid and the switching digital logic which represents a resistive path to ground. Therefore the transition time for any layer voltage is dictated by the RC time constant due to the capacitance of the stack layers and the resistive discharge path to ground. Figure 4.9 shows that in the case of the test chip presented in this chapter, it takes ∼300ns for the voltage to transition from 0.9V to the worst-case minimum voltage shown in Figures 4.9(c)-(e). Adding more decoupling capacitance to each layer will decrease the rate of voltage transition, but for static activity mismatch, it ultimately cannot alleviate voltage droop. Typically, inductance of on-chip power grid is small enough that inductive effects are negligible for inter-layer noise in voltage stacking.

In voltage stacking, all layer voltages must add up to V_{IN} and, therefore, all layer voltages are inherently interdependent. Box plots of Figures 4.9(b)-(e) show that voltage droop in V_{SL2} results in increase in other layer voltages. To present a complete picture of how activity mismatch affects layer voltages, Figure 4.10 presents waveforms of V_{SL2} for a variety of other activity mismatch conditions. Box plots of all four stack layer voltages are also presented for completeness. As discussed previously, a comparison of Figures 4.10(a) and (b) shows that voltage droop reduces when more layers execute higher activity workloads. Other scenarios, in Figure 4.10(c)-(e), lead to increases in V_{SL2} due to higher activity workloads running on cores in other layers. We can infer the general behavior of voltage noise in all layers from the measured noise profiles presented in Figure 4.9 and Figure 4.10.

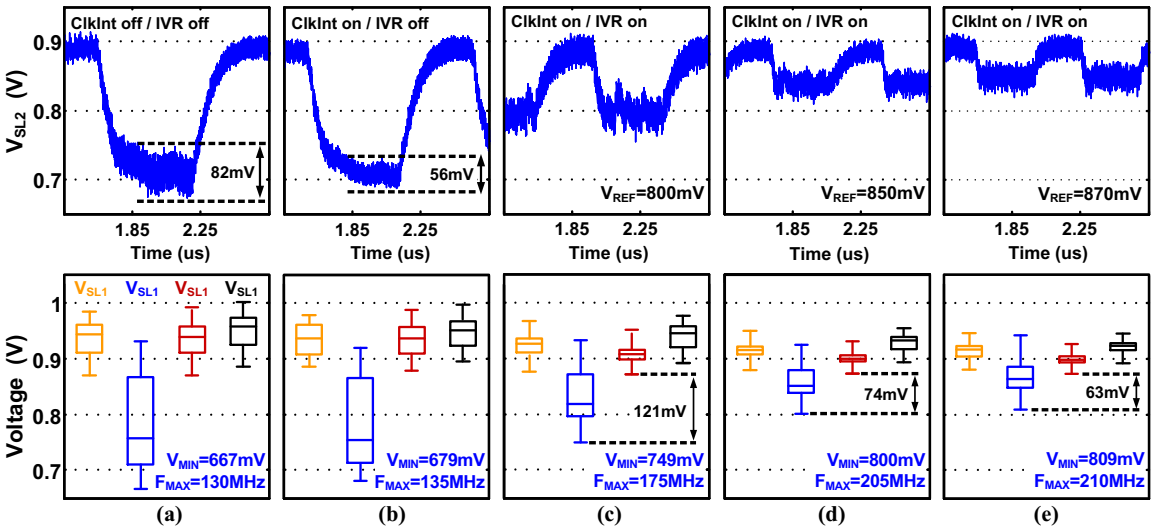


Figure 4.11: Measured transient waveforms of V_{SL2} and boxplots of V_{SL1}-V_{SL4} for unbalanced workload scenario running FFClk at various test chip settings: (a) ClkInt off, IVR off (b) ClkInt on, IVR off (c) ClkInt on, IVR on (V_{REF}=800mV) (d) ClkInt on, IVR on (V_{REF}=850mV) (e) ClkInt on, IVR on (V_{REF}=870mV)

4.3.3 Mitigating Voltage Noise for Fixed-Frequency Clocking Operation using an IVR

The worst-case voltage noise observed in Figure 4.9(c)-(e) is especially problematic for global fixed-frequency operation (FFClk), because the global clock frequency must be chosen to ensure correct operation for all cores at the minimum operating voltage (V_{MIN}) under worst-case voltage noise conditions, requiring large voltage margins. Moreover, there are additional constraints on V_{MIN}. For example, it must be high enough to ensure reliable operation of voltage-sensitive circuits such as on-chip SRAMs, which can exhibit SRAM read/write failures under low voltage conditions. Consequently, the test chip includes an IVR to provide certain V_{MIN} guarantees.

To demonstrate the IVR's abilty to reduce inter-layer voltage noise, let us begin with the worst-case noise scenario previously seen in Figure 4.9(c). A close-up of V_{SL2},

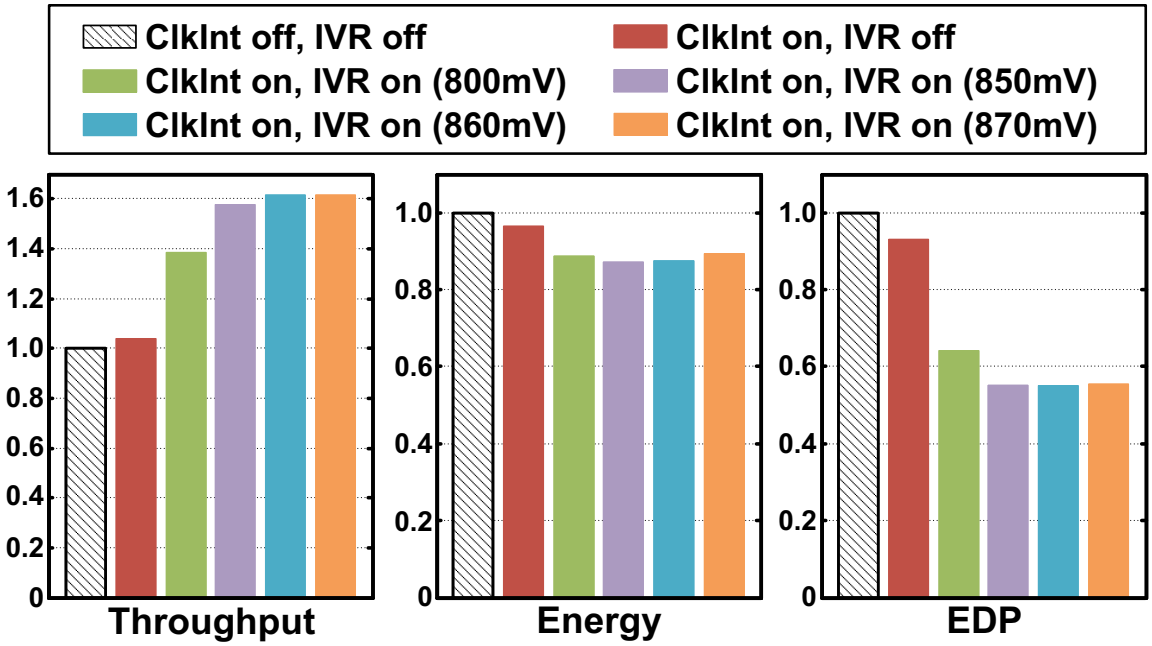


Figure 4.12: Normalized system throughput, energy, and EDP of unbalanced testing scenarios running FFClk under various test chip settings.

in Figure 4.11(a), shows the minimum voltage observed over a 1ms execution window is 667mV. Based on memory BIST tests, all memories were functional down to 650mV and, therefore, the test scenario avoids issues due to SRAM instability. However, core performance degradation is a concern. To accommodate the worst-case voltage noise, voltage margin in excess of 230mV is required, limiting the maximum operating frequency (F_{MAX}) to only 130MHz. Figure 4.11(b) first demonstrates the benefits of clock-phase interleaving (ClkInt). Although within-cycle curent smoothing reduces voltage ripple, the minimum voltage is still less than 700mV. In contrast, also turning on the IVR provides much more pronounced benefits, as shown in Figures 4.11(c)-(e) for different V_{REF} settings. These results confirm the IVR can maintain V_{SL2} around V_{REF} and provide V_{MIN} guarantees, which then allows the cores to operate at a higher F_{MAX}. However, it is important to note the IVR cannot regulate the

layer voltages to exactly V_{NOM}. Hence, there are diminishing returns of setting V_{REF} too aggressively. Furthermore, Figures 4.11(c)-(e) show that cores in SL1, SL3, and SL4 must operate at a higher power for the same performance level, due to the extra voltage margin caused by the elevated stack layer voltages, which degrades the overall system energy-efficiency.

To quantify the benefits gained from clock-phase interleaving and turning on the IVR, Figure 4.12 presents the aggregate throughput, energy, and EDP of the test chip settings shown in Figure 4.11, normalized to the baseline case when none of the features are turned on (shown in Figure 4.11(a)). The results show that the V_{MIN} improvements gained from turning on the IVR mostly improves average throughput enhancements, which then translates to large EDP reductions. Higher V_{MIN} also reduces voltage margins to also provide energy improvements despite incurring IVR energy overheads. However, aggressively increasing V_{REF} exacerbates IVR losses without improving V_{MIN} or throughput.

4.3.4 Adaptive Frequency Clocking with IVR

To further reduce voltage margin and thereby further improve system performance and energy-efficiency, the test chip implements adaptive frequency clocking (AFClk). The clock frequency of the DCRO in each layer tracks stack layer voltage fluctuations to independently minimize voltage margins per layer. Consequently, AFClk does not penalize stack layers for voltage noise in other layers. Fig 4.13 presents the voltage waveform of V_{SL2}, box plots of all layer voltages, and the per-layer DCRO clock frequencies for the same unbalanced scenario presented before in Figure 4.11 with ClkInt

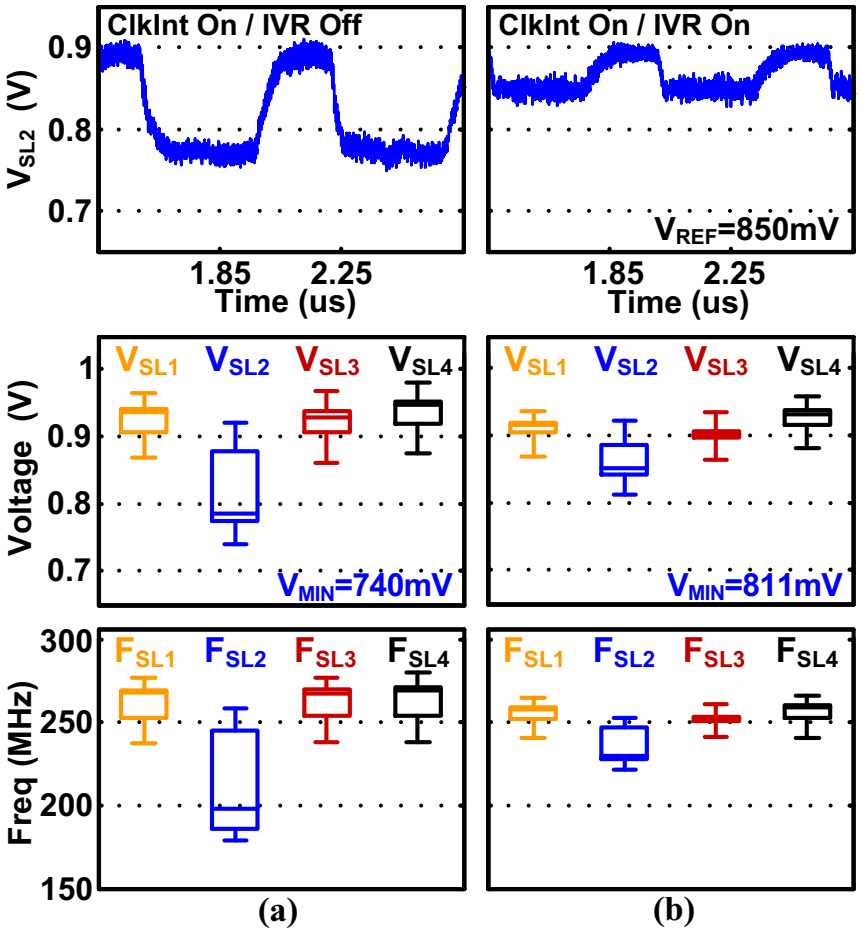


Figure 4.13: Measured transient waveforms of V_{SL2} (top), box plots of V_{SL1}-V_{SL4} (middle) and box plots of per-layer DCRO frequency distribution (bottom) for unbalanced workload scenario running AFClk at two different test chip settings: (a) ClkInt on, IVR off. (b) ClkInt on, IVR on (V_{REF}=850mV).

turned on. With the IVR off (Figure 4.13(a)), AFClk results in smaller voltage droop (higher V_{MIN}) compared to FFClk operation. This is because per-layer dynamic current consumption has a stronger dependence on voltage compared to FFClk since AFClk frequency depends on voltage. Moreover, the spread in clock frequencies, seen in Figure 4.13(a), demonstrates that rather than being limited to the worst-case frequency set by noise in V_{SL2}, cores in SL1, SL3, and SL4 ran at higher frequencies set

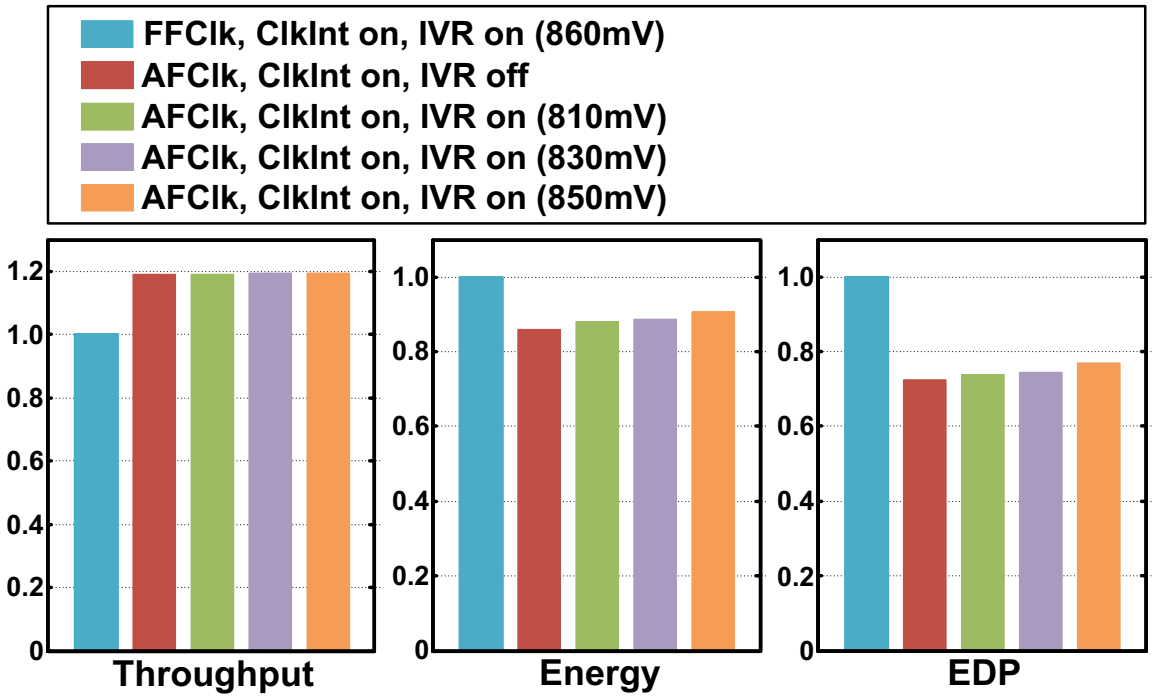


Figure 4.14: Normalized aggregate throughput, energy, and EDP comparison of adaptive frequency clocking operation (AFClk) versus fixed-frequency operation (FFClk) for various test chip settings under worst-case noise condition. IVR V_{REF} is noted in parenthesis for the cases when IVR is on.

by the DCRO in each layer operating at higher voltages, thereby improving overall system performance and energy-efficiency. Figure 4.13(b) demonstrates AFClk operation with the IVR turned ON and V_{REF} set at 850mV. The results show that AFClk allows the cores to operate with minimal voltage margin, while the IVR maintains a V_{MIN} (>810mV in this case). This provides a minimum performance guarantee for all layers and protection against potential reliability issues (e.g., SRAM instability).

Putting the measured results together, Figure 4.14 presents the resulting aggregate throughput, energy, and EDP for AFClk. The results are normalized to that of FFClk operation with the IVR ON and V_{REF} set to 860mV, shown in Figure 4.12, which is the highest performance and energy-efficiency V_{REF} setting for FFClk. Since cores

in each layer operate at different voltage/frequency values for AFClk due to voltage noise, we calculate the system energy as the sum of total energy consumed by all cores to complete the same amount of work:

$$E = \sum_{i=\mathrm{SL1}}^{\mathrm{SL4}} P_i t_i + P_{\mathrm{IVR}} t_{\mathrm{SL2}} \tag{4.2}$$

where $P_{\mathrm{SL}\{1:4\}}$ is the power consumption of the entire layer that includes all four cores and the DCRO. IVR energy overhead is calculated by multiplying IVR power loss ($\mathrm{P_{IVR}}$) with the time it takes for the cores in SL2 to complete the task, because the IVR is supplying extra current only to the cores in SL2 in this workload scenario. EDP is also calculated in a similar fashion as:

$$EDP = \sum_{i=\mathrm{SL1}}^{\mathrm{SL4}} P_i t_i^2 + P_{\mathrm{IVR}} t_{\mathrm{SL2}}^2 \tag{4.3}$$

The results show that even when compared against the best-case FFClk scenario, AFClk provides significant improvements in throughput, energy, and EDP due to reduction in voltage margin for all cores. Notice the aggregate throughput of the system for adaptive frequency clocking stay relatively constant independent of whether the IVR is ON or not. This is because in the high-throughput scenario with all 16 cores active, loss of performance (frequency) in any layer is compensated by higher frequency in other layers. Overall, AFClk achieves reduction in energy and EDP compared to the best-case FFClk, even when accounting for extra power consumed by the DCRO and the IVR energy overhead. Setting higher IVR $\mathrm{V_{REF}}$ for higher $\mathrm{V_{MIN}}$ translates to increases in energy and EDP, as expected, due to higher IVR energy overhead.

The results shown in Figures 4.13 and 4.14 clearly demonstrate the advantages of AFClk over FFClk for the voltage-stacked system presented in this chapter. For all subsequent test results, we use AFClk as the default clocking mode of operation to achieve best possible performance and efficiency for the system.

4.4 System-wide Power Delivery Efficiency

As shown in the energy comparisons presented previously, voltage regulation using the IVR comes with an energy overhead. Typically, IVR losses are quantified by the IVR efficiency. For conventional systems in which the IVR provides the entirety of the power consumed by the system, IVR efficiency represents the overall on-chip power delivery efficiency. However, voltage-stacked systems are fundamentally different in that they require the IVR to only supply the load current necessary to compensate for inter-layer activity mismatch. Therefore, IVR losses only apply to the mismatch-related power, while stack current that is common to all layers, such as leakage, is recycled efficiently through the stack. For this reason, IVR efficiency does not sufficiently capture the overall power delivery efficiency in voltage stacking.

To properly quantify power delivery efficiency in voltage-stacked systems, we define the metric *system-wide power delivery efficiency* (η_{SYS}) as:

$$\eta_{\mathrm{SYS}} = \frac{Power\ consumed\ by\ all\ cores}{Total\ power\ delivered\ to\ test\ chip}$$

Figure 4.15 presents the measured IVR efficiency (η_{IVR}) and the system-wide power delivery efficiency (η_{SYS}) when only the cores in SL2 consume extra power. Figure 4.15(a) illustrates the experimental setup. With all cores in SL1, SL3, and

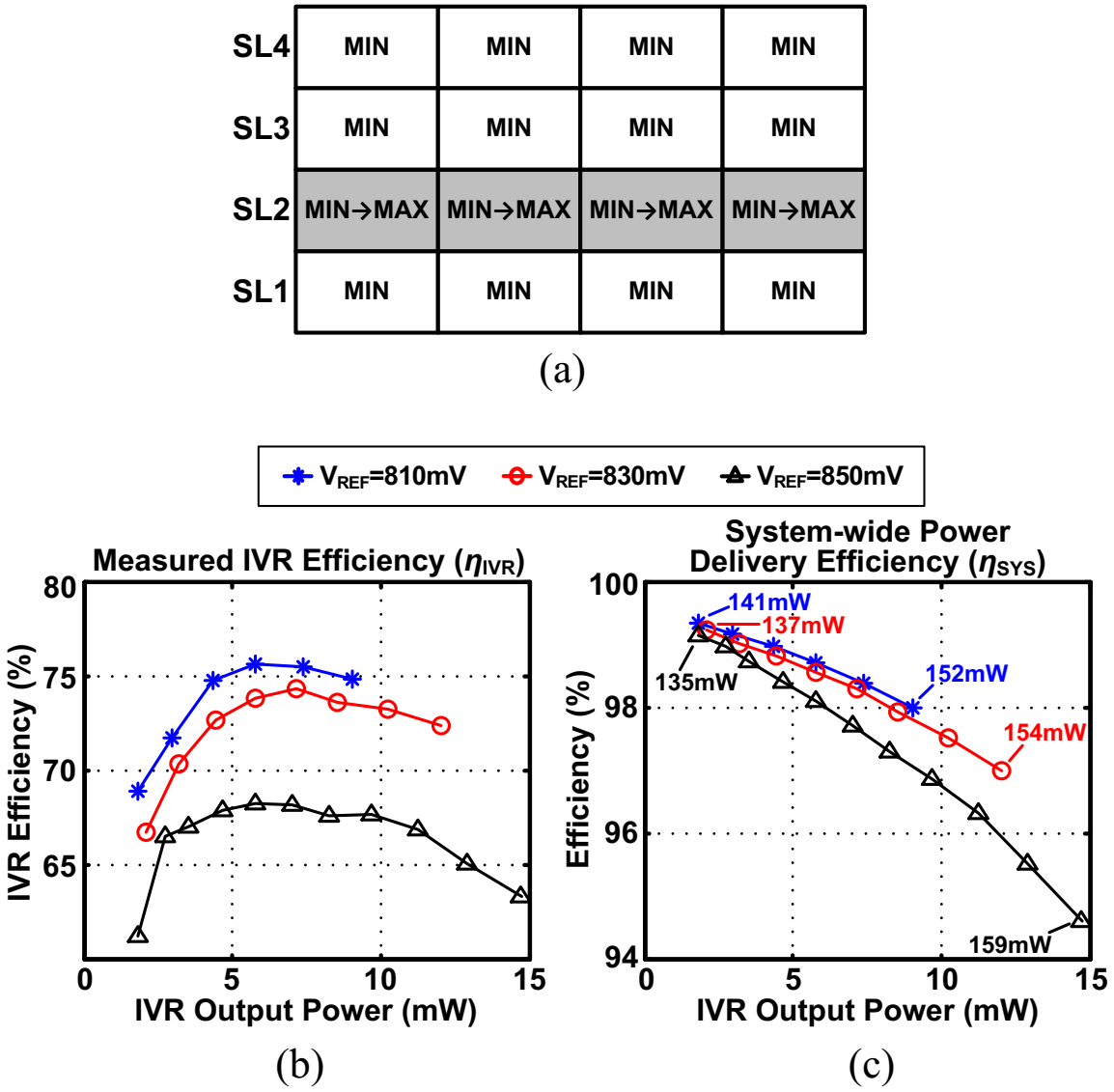


Figure 4.15: (a) Block diagram illustrating the test setup to characterize IVR and system-wide power delivery efficiency. (b) Plot of measured IVR efficiency when delivering power to SL2 for three V_{REF} levels. (c) The corresponding system-wide power delivery efficiency measurements with annotations of the total power consumed by the test chip for the first and the last values of each V_{REF}.

SL4 executing the *MIN-CONT* micro-benchmark, power consumption of the four cores in SL2 was swept from minimum to maximum by executing ISA characterization micro-benchmarks used for instruction power profiling. Figure 4.15(b) and (c) plot

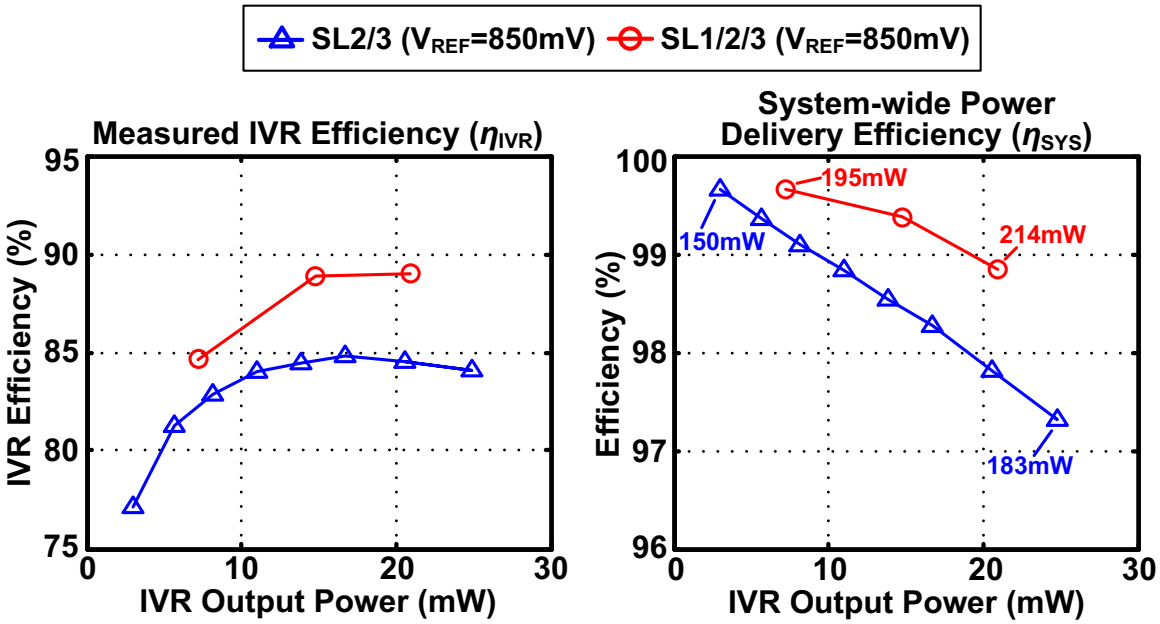


Figure 4.16: Measured IVR efficiency (left) and corresponding system-wide power delivery efficiency (right) for two workload scenarios with VtextsubscriptREF set to 850mV.

the η_{IVR} and η_{SYS} of the test chip, respectively, for three IVR V_{REF} settings. Total of 20 micro-benchmarks were run on the cores in SL2, from *MIN-CONT* (minimum power) up to *MAX-CONT* (maximum power). Progressively higher power micro-benchmarks leads to larger voltage droop that then requires more IVR activity and output power. The results show that higher V_{REF} translates to wider operating range but lower η_{IVR}, because the IVR must work harder to drive the layer voltage to higher levels. Consequently, η_{SYS} also degrades for higher V_{REF}. However, more importantly, Figure 4.15(c) confirms the IVR losses only apply to a small fraction of total power consumed by all of the cores and η_{SYS} is greater than 94% across all of the experiments.

Spreading out higher activity across multiple layers leads to less inter-layer volt-

age droop than concentrating all of the higher activity to a single layer. Moreover, the symmetric-ladder switched-capacitor IVR implemented on the test chip is more efficient at supplying power to multiple layers simultaneously than to a single layer. Due to its topology, delivering power to multiple layers reduces the total amount of charge flowing through the flying capacitors and power switches, reducing the losses [34, 61]. To demonstrate this aspect of the system, Figure 4.16 plots the measured IVR efficiency and system-wide power delivery efficiency when the IVR delivers power to multiple layers simultaneously. This experiment is similar to the prior experiment run for Figure 4.15, but with the power swept from minimum to maximum in two layers (SL2 and SL3) and three layers (SL1,SL2 and SL3). IVR V_{REF} is set to 850mV for both cases. Overall, both the η_{IVR} and η_{SYS} improves when the IVR must deliver power to multiple layers, even when the total test chip power is higher than the conditions corresponding to the results in Figure 4.15.

4.5 Workload and Core Allocation

The test results so far have demonstrated the IVR and adaptive frequency clocking to be effective hardware solutions to improve system throughput and energy-efficiency and to provide the necessary minimum voltage guarantees required by the voltage-stacked system. We now take a step back to explore a software solution that relies on intelligent inter-layer workload balancing to further improve overall efficiency. This section explores the impact of workload and core allocation in multicore voltage-stacked systems.

Recall the worst-case workload scenario presented in Figure 4.15, where all cores in

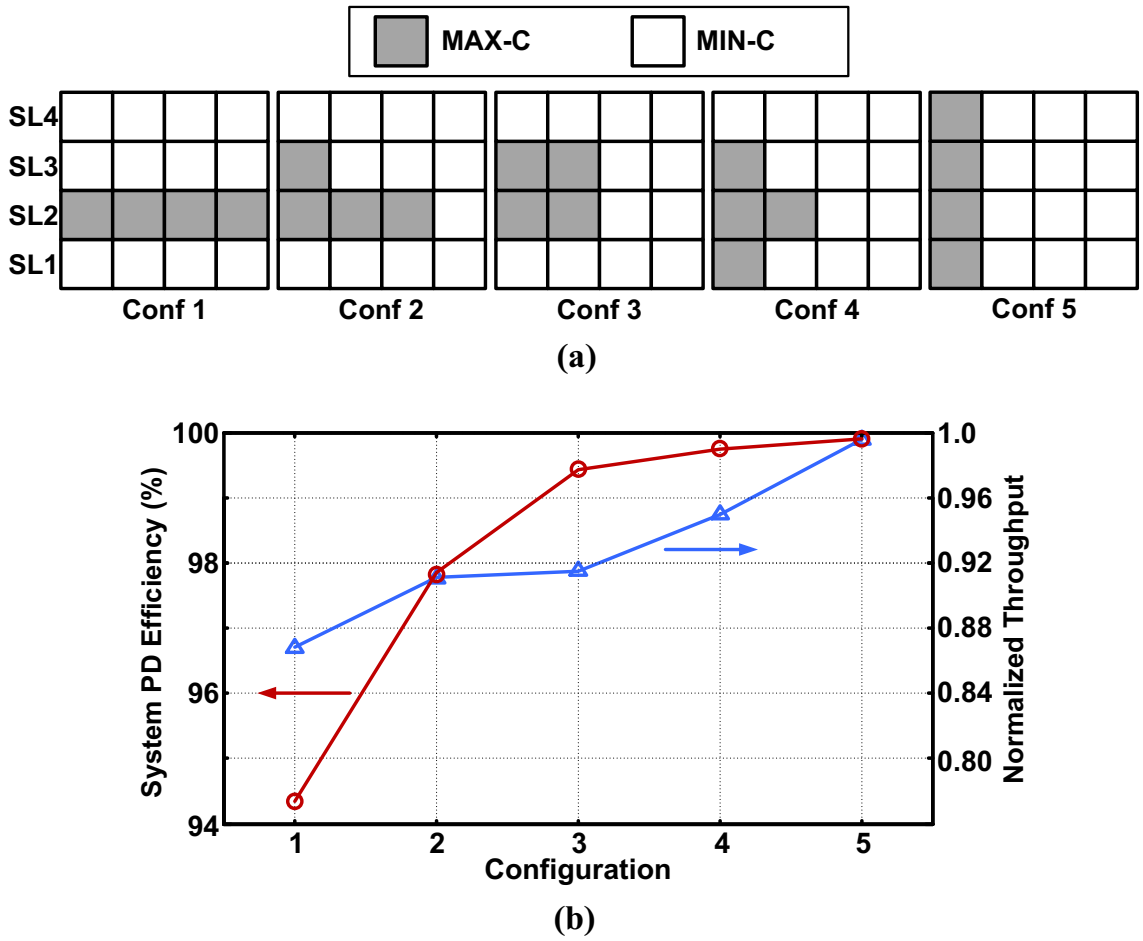


Figure 4.17: (a) Five possible workload allocation scenarios with the same set of workloads. (b) Normalized average per-core throughput for the four cores running MAX-CONT and IVR loss.

SL2 execute the *MAX-CONT* micro-benchmark while all other cores run *MIN-CONT*. Figure 4.17(a) presents five different workload allocation scenarios that are possible for the same set of workloads running on the 16 cores. The workload allocation becomes progressively more balanced going from configuration 1 to 5. To show the benefits of a more balanced workload allocation, Figure 4.17(b) plots the IVR loss for all configurations as well as the normalized average throughput of the four cores that are running *MAX-CONT*. Measurement results show that better balancing workloads

Table 4.3: Kernels used for testing

Benchmark	Description
AES	AES encryption
GEMM/BLOCKED	Matrix multiplication
GEMM/NCUBED	Matrix multiplication
KMP	String matching
MD/GRID	Molecular Dynamics
MD/KNN	Molecular dynamics
NW	DNA alignment
REDUCTION	Sum reduction operation
SCAN	Parallel prefix sum
SHA	Cryptography hash function
SORT/MERGE	Sorting
SORT/RADIX	Sorting
STENCIL/STENCIL2D	Sparse Matrix/vector multiplication
STENCIL/STENCIL3D	Sparse Matrix/vector multiplication
TRIAD	Vector dot product
VITERBI	Hidden Markov model estimation

Table 4.4: Workload scenarios for efficiency tests

Workload Scenario	Description
RAND-LYR	Random kernel allocated per layer
RAND-CORE	Random kernel allocated per core
BALANCED	All cores running the same kernel

across the layers leads to lower voltage noise, higher system-wide power delivery efficiency, and higher throughput.

To further expand the experiment beyond the worst-case scenario and demonstrate the effect of workload allocation on the efficiency of the system running real kernels, we

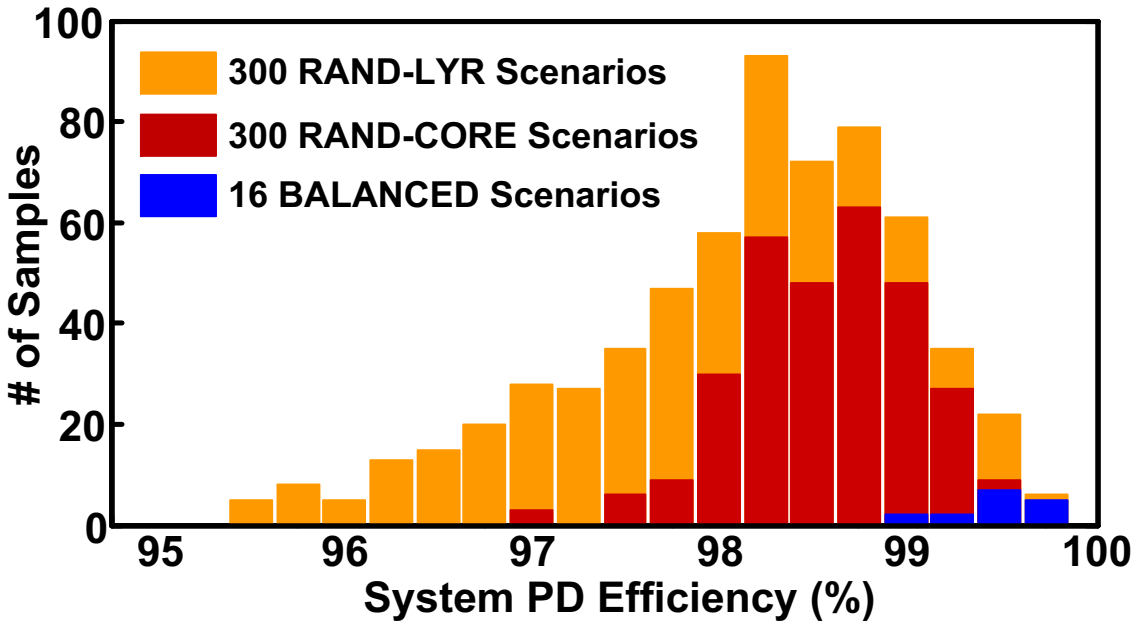


Figure 4.18: Histogram of system-wide power delivery efficiency for 616 workload scenarios described in Table 4.4.

chose 16 kernels to run on the cores from the *Machsuite* benchmark suite [49]. These kernels cover a broad range of applications including matrix multiplication, sort, string matching, AES encryption, molecular dynamics, and Viterbi algorithm, and offer a diverse mixture of workload behaviors (compute-intensive, memory-intensive...). The kernels are listed in Table 4.3. To analyze the effects of different workload behaviors on efficiency, we generate three different groups of workload scenarios using the 16 kernels to run on the system, as summarized in Table 4.4. Figure 4.18 plots the resulting histogram of system-wide power delivery efficiency, which is greater than 95% for all workload scenarios. *RAND-LYR* scenarios exhibit the worst efficiency, mainly because they have the worst inter-layer activity mismatch due to cores in each layer running the same kernel, effectively amplifying the mismatch. When each

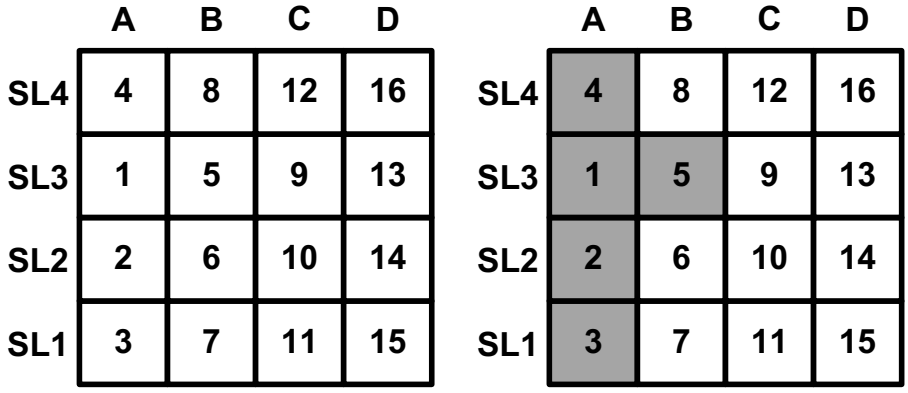


Figure 4.19: Block diagram illustrating the order in which the cores are activated when not all 16 cores are active (left), and example of the active cores (shaded) for configuration 5 when only five cores are active and all other cores are powered down (right).

of the 16 cores run one of 16 randomly chosen kernels (*RAND-CORE*), there is an averaging effect on the per-layer core activities, which ends up reducing overall voltage noise and IVR losses. Finally, with the *BALANCED* scenarios, efficiency is 99% or better, because the IVR rarely turns on.

In addition to workload differences, stack layers may also suffer imbalances due to coarser-grained core inactivity. Further expanding on the concept of intelligent workload allocation, we deduce that for any number of active cores, there is an optimum allocation that evenly distributes work to minimize inter-layer activity mismatch and maximize efficient current recycling through the stack. Figure 4.19(a) illustrates the optimum order that work should be allocated across the 16 cores. For example, only the five cores labeled 1 through to 5 will be active for configuration 5, as shown in Figure 4.19(b), while the rest are inactive. For configuration 6, Core B-SL2 (labeled 6) will be activated in addition to the cores in configuration 5. As more cores are turned on, the principle is to activate cores vertically along a column rather than hor-

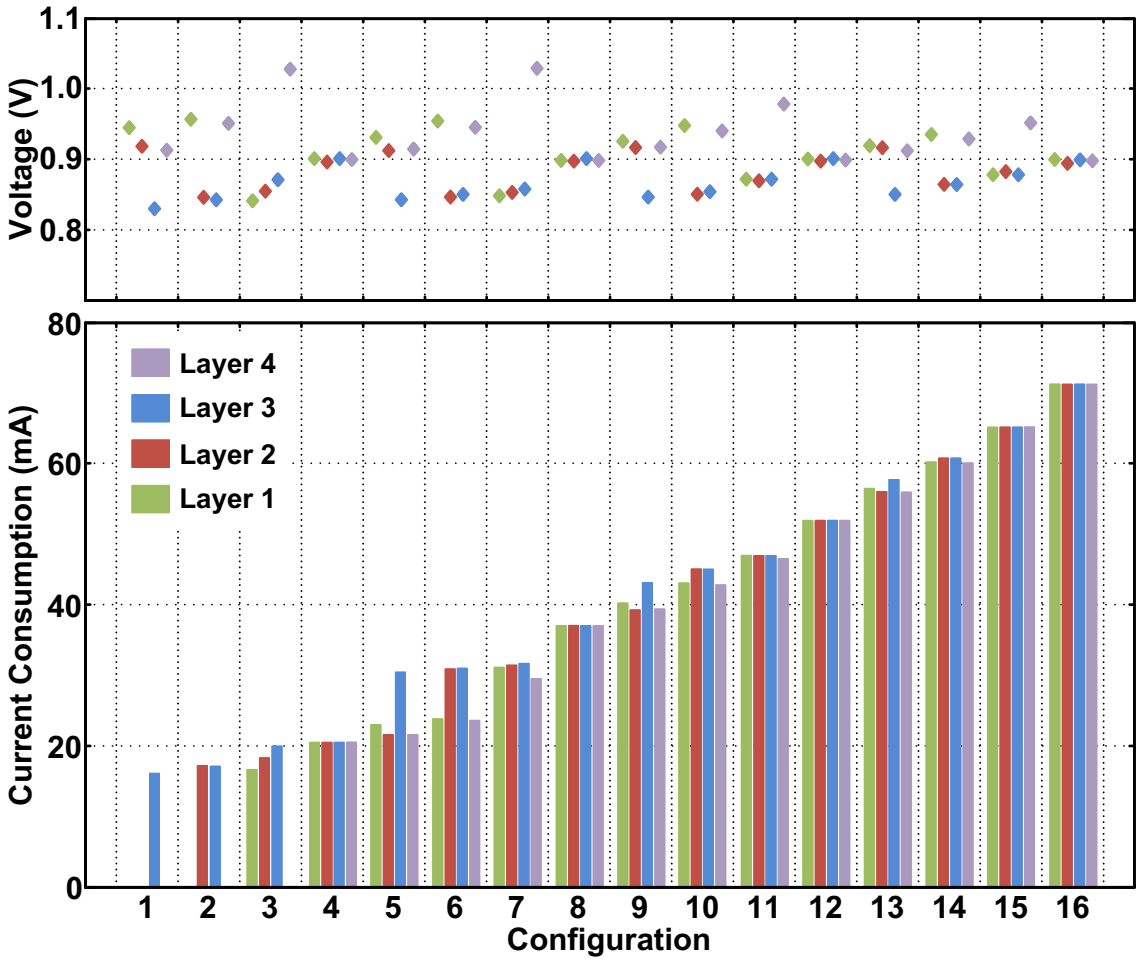


Figure 4.20: Average per-layer operating voltage and current consumption for all 16 configurations configuration.

izontally along the same layer. For voltage stacking, this achieves the most balanced core allocation possible for any number of active cores. Note that for each column, the cores in SL3 and SL2 are activated before turning on the cores in SL1 and SL4. This is to take advantage of the fact that the IVR is more efficient in delivering power to the middle two layers [34, 61]. Allocating cores this way ensures that the maximum activity mismatch between any layers does not exceed the activity of a single core, preemptively capping the power that the IVR has to process.

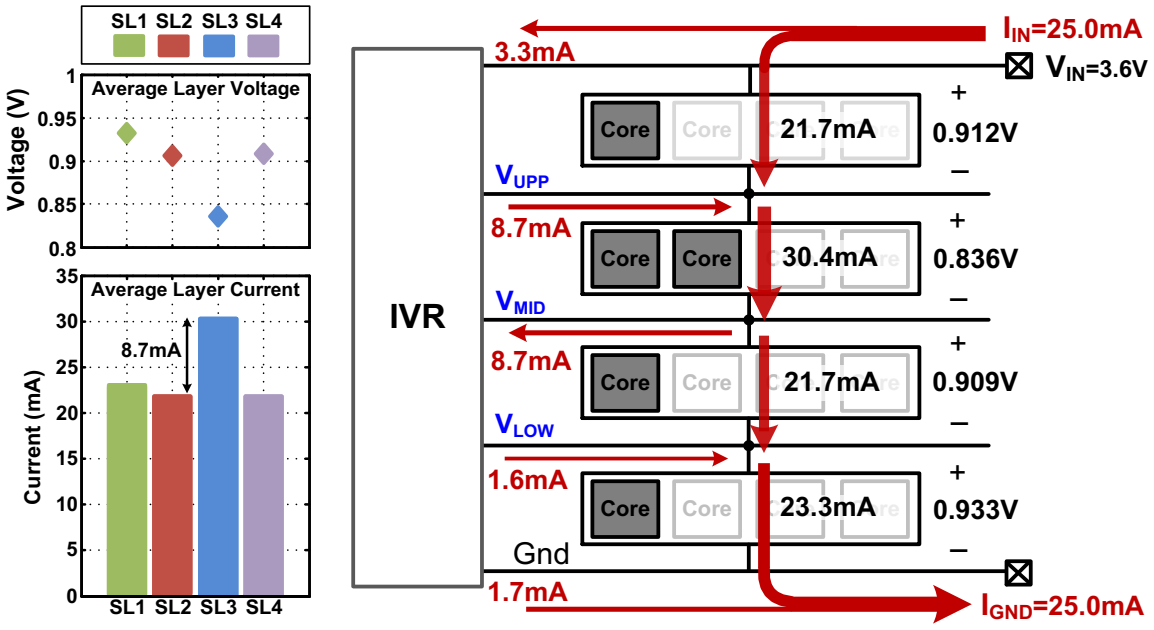


Figure 4.21: Block diagram illustrating the current flow through the test chip for configuration 6 in Figure 4.19.

We can now run an experiment to understand the effects of coarse-grained core activity/inactivity and workload allocation. First, all inactive cores are fully powered OFF so that leakage current consumption of these cores is negligible. Second, all cores execute the *MAX-CONT* micro-benchmark. Figure 4.20 presents the average operating voltage and current consumption for each stack layer for all 16 configurations. Figure 4.21 presents an illustration of the current flow through the stack layers for configuration 5 presented in Figure 4.20. Average per-layer voltage and current are also shown for clarity. Given the extra load in SL3, V_{SL3} droop below V_{REF} and the IVR maintains the minimum V_{SL3} close to V_{REF} (set to 850mV). The average current consumption of 21.7mA common to all stack layers is recycled through the stack, while the IVR only provides the additional 8.7mA of current that is required to support the two cores in SL3, and the additional 1.6mA of current that is required

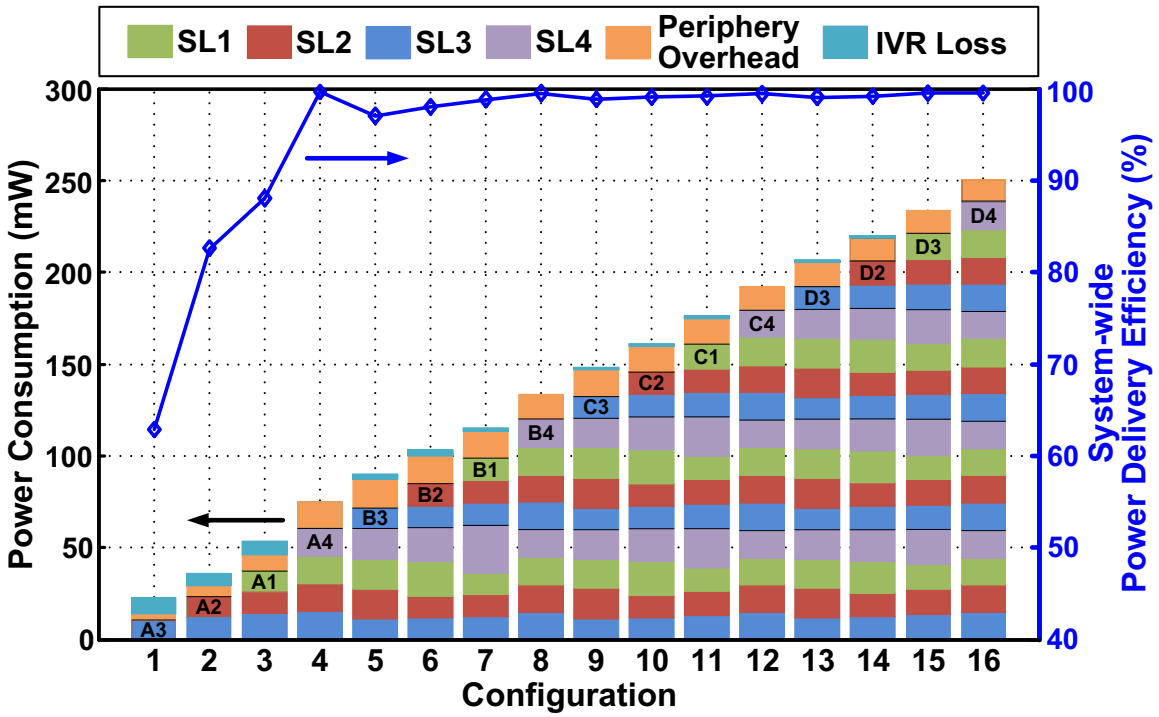


Figure 4.22: Power consumption of the test chip and the corresponding system-wide power delivery efficiency for the 16 different core activation configurations. For each configuration, the additional core that is activated in addition to the previous configuration is marked with the coordinates as described in Figure 4.19.

to support the core in SL1 operating at an elevated voltage. With the IVR delivering additional current to support the extra load in SL3, the distribution of V_{SL1}, V_{SL2}, and V_{SL4} is determined by the static behavior of the IVR, which dictates that different amounts of charge must flow through the flying capacitors of the symmetric ladder, resulting in higher V_{SL1} than V_{SL2} and V_{SL4}.

The average per-layer voltage and current presented in Figure 4.20 provides insight into the amount of power that the IVR has to process for different numbers of active cores under this core allocation scheme. For configurations 1 to 3, where one or more layers are powered off, no current can be recycled through the stack. Therefore, the IVR must supply all of the power consumed by the active cores. Beyond configuration

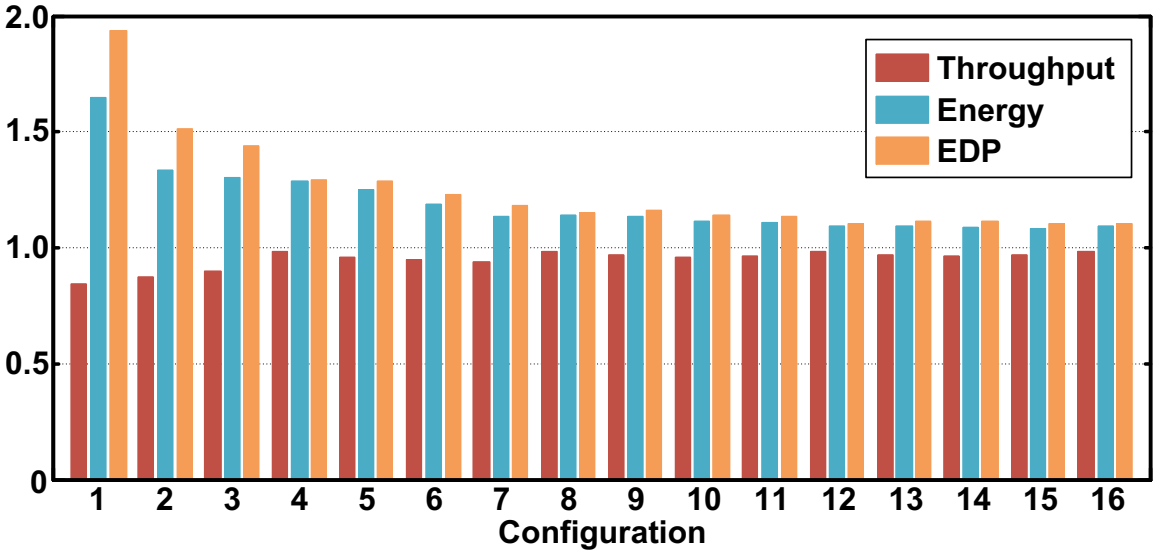


Figure 4.23: Normalized average per-core throughput, energy and EDP for the 16 different core activation configurations.

4, however, when at least a full column of cores are active, current consumption common to all layers are recycled through the stack while the IVR only supplies the additional current required to support the higher activity due to additional active cores in the stack layers. Since all cores are running the same micro-benchmark in this experiment, configurations 4,8,12 and 16 are balanced scenarios and the layer voltages evenly subdivide V_{IN} to 0.9V. A comparison of results for configurations 1, 5, 9, and 13 in Figure 4.20 confirms the earlier observation in Chapter 3 that higher overall current flow through the stack alleviates voltage noise for the same inter-layer activity mismatch.

Figure 4.22 presents a stacked bar chart of the power consumption for each configuration as well as the corresponding system-wide power delivery efficiency. IVR loss and periphery overhead, which includes the DCRO, clock distribution, and all other uncore power such as layer shifter power, are shown separately. This figure

provides a measure of the IVR losses incurred by the system in relation to the overall power delivered. For configurations 1 through 3, where the IVR supplies all of the power consumed by the cores, the IVR loss is substantial, and the system-wide power delivery efficiency equals the IVR efficiency. As more cores are activated, however, inter-layer charge recycling increases and IVR losses reduce. Overall system-wide power delivery efficiency exceeds 95% for all configurations when four or more cores are active.

Finally, Figure 4.23 presents the average per-core throughput, energy, and EDP for all 16 configurations. The results are normalized to an ideal baseline scenario where all layer voltages are losslessly driven to V_{NOM}. For 1 to 3 active cores, energy and EDP penalties are high due to large IVR loss and DCRO overhead, coupled with the loss of performance. As more cores turn on, the overall energy and EDP of the system improves due to decreasing IVR loss, amortized per-core DCRO overhead, and more balanced voltage distribution, leading to smaller loss in throughput.

Chapter 5

Coupled Communication Circuits for Efficient Inter-layer Communication

In addition to robust noise mitigation techniques shown in the previous chapter, voltage stacked systems require efficient on-die signaling circuitry to enable communication between series-connected stack layers that do not share common voltage reference planes. Direct communication between different stack layers requires isolation to avoid dc current paths between stacked voltage rails, while direct communication between stacked voltage domains not in close physical proximity requires efficient repeater-less long-distance signaling, because repeater placement may not be possible en route. Finally, deviations of internal voltage rails due to inter-layer power consumption mismatch, as shown in Chapter 4, necessitate reliable communication across a wide range of stack layer voltages.

Prior works demonstrated capacitively-coupled full-swing layer shifters with cross-coupled inverter latch receivers to shift signals across stack layer boundaries using MOS capacitor [46]. However,the use of MOS capacitors for coupling prohibit direct communication between non-adjacent stack layers due to device overstress issues, requiring relay communication which incurs additional energy overhead, and these layer shifters also suffer from start-up issues that degrade reliability. Even when implemented using metal-oxide-metal (MoM) capacitors, as in the 16-core test chip presented in Chapter 4 (shown in Figure 4.8), these designs incur large capacitor and driver size overhead, required to flip the cross-coupled inverter latch, especially when communicating between stack layers not in close physical proximity.

To address these issues, this chapter presents layer shifting communication circuits that send pulsed signals across capacitively- and inductively-coupled passive elements to a common differential hysteresis receiver. Capacitively- or inductively-coupled pulse-mode communication was previously demonstrated for efficient repeater-less long-distance on-chip communication and for efficient inter-chip communication [22, 56, 40, 39, 71, 36, 53]. Proposed layer shifters leverage the efficiency and high performance of low-swing pulse-mode signaling with voltage isolating passive elements to enable highly efficient communication between the stacked voltage domains, while also enabling efficient broadcasting of signals from one to multiple stack layers.

Section 5.1 presents an overview of a test chip prototype that implements capacitively- and inductively-coupled communication schemes for inter-layer signaling for a 3-way voltage stacked system that implements 1-to-3 layer broadcast communication in addition to 1-to-1 communication between stack layers. Sections 5.2 and 5.3 provides

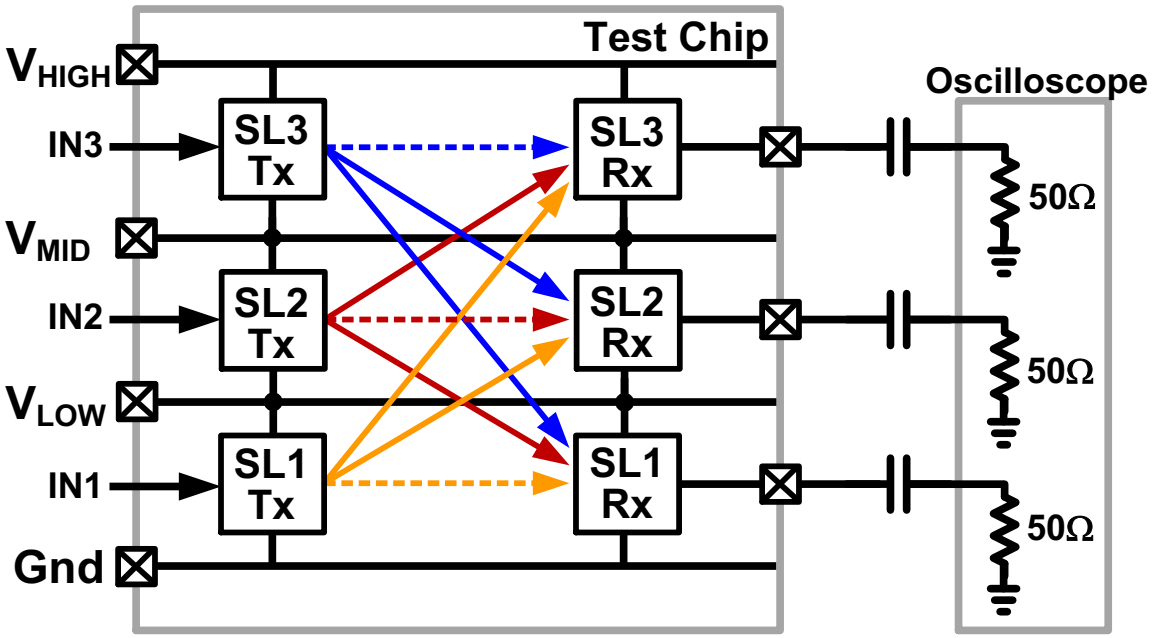


Figure 5.1: Block diagram of voltage-stacked inter-layer communication test chip prototype and test environment.

a detailed description of both communication schemes implemented on the test chip. The experimental results presented in Section 5.5 demonstrates efficient communication at 3.4Gb/s.

5.1 Overview of 3-layer Communication Test Chip Prototype

Figure 5.1 presents an overview of the test chip fabricated in TSMCs 40G process. The test chip implements a voltage stacked inter-layer communication system with three stacked layers. External voltage sources supply V_{HIGH}, V_{MID}, V_{LOW} with 2.7V, 1.8V, 0.9V, respectively, with each stack layer (SL1, SL2, SL3) operating nominally

at 0.9V. The stack layers are isolated from body bias effects using triple-wells. The test chip implements capacitively- and inductively-coupled transmitter-receiver pairs across stack layer boundaries to demonstrate layer shifting communication. Input signals to the layer shifters are driven from off-chip, while the outputs from each stack layer are driven through an off-chip DC blocking capacitor to the oscilloscope to avoid DC current paths. The test chip implements all possible 1-to-1 inter-layer paths between the three stack layers, including within-layer communication paths (dotted lines in Figure 5.1), for both coupled communication schemes. The test chip also implements 1-to-3 signal paths that combine one transmitter from a single stack layer with three receivers one from each layer to demonstrate broadcast communication from one to multiple layers. Within-layer communication paths facilitate the demonstration of 1-to-3 inter-layer signal broadcast without the complexity of adding an extra stack layer in the test chip.

5.2 Capacitively-Coupled Layer Shifter

Figure 5.2 (top) presents a block diagram of the capacitively-coupled layer shifter implemented on-chip that drives signals across stack layer boundaries using series capacitors built with MoM caps. MoM caps provide effective isolation of the stacked Tx and Rx voltage domains, within the inter-metal dielectric breakdown limit. The Tx driver generates differential return-to-zero (RZ) pulses on the wires by driving the Tx caps C_{C1} and C_{C2} ($=0.3C_{C1}$) in a push-pull manner. Pulses are only generated on data transitions and the pulse width is determined by the delay cell, whose delay is 150ps at 0.9V. A clock-less hysteresis latch receiver shown in Figure 5.2 (lower right)

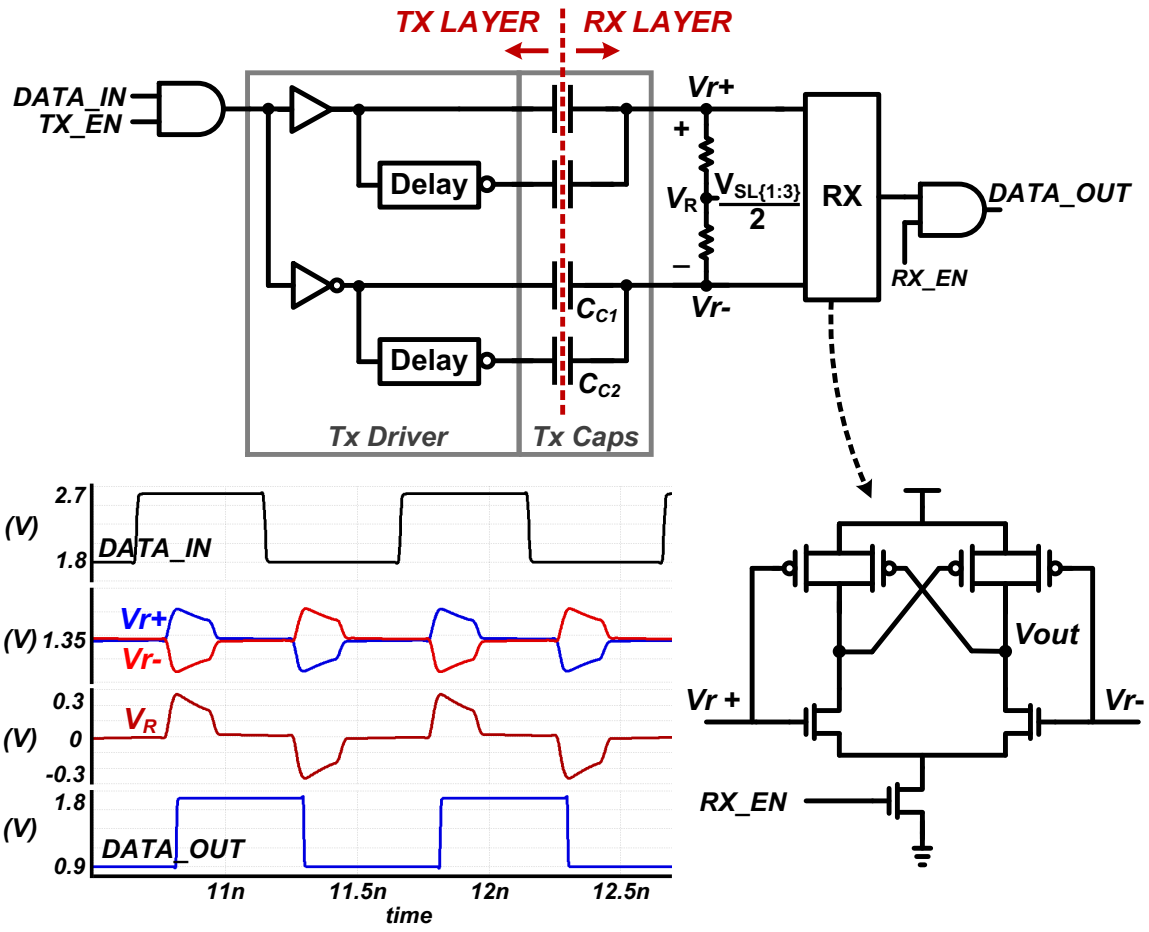


Figure 5.2: Block diagram of capacitively-coupled layer shifter (top), schematic of the hysteresis latch receiver (lower right) and simulation waveforms (lower left).

senses the resulting differential pulse V_R to recover the transmit data. Receiver hysteresis is 75mV at 0.9V. The receiver input biases are generated on-chip using resistor dividers. Hysteresis receivers are simpler than the clocked receivers often used in inter-chip pulse-mode signaling and they consume no energy on idle inputs [56]. Simulation waveforms presented in Figure 5.1 (lower left) demonstrate data transmission from layer SL3 to SL2. Compared to the full-swing layer shifters in [46] and Chapter 4, this approach requires smaller capacitors due to low-swing pulse-mode signaling, reducing the driver load. The proposed layer shifter requires 10fF of series

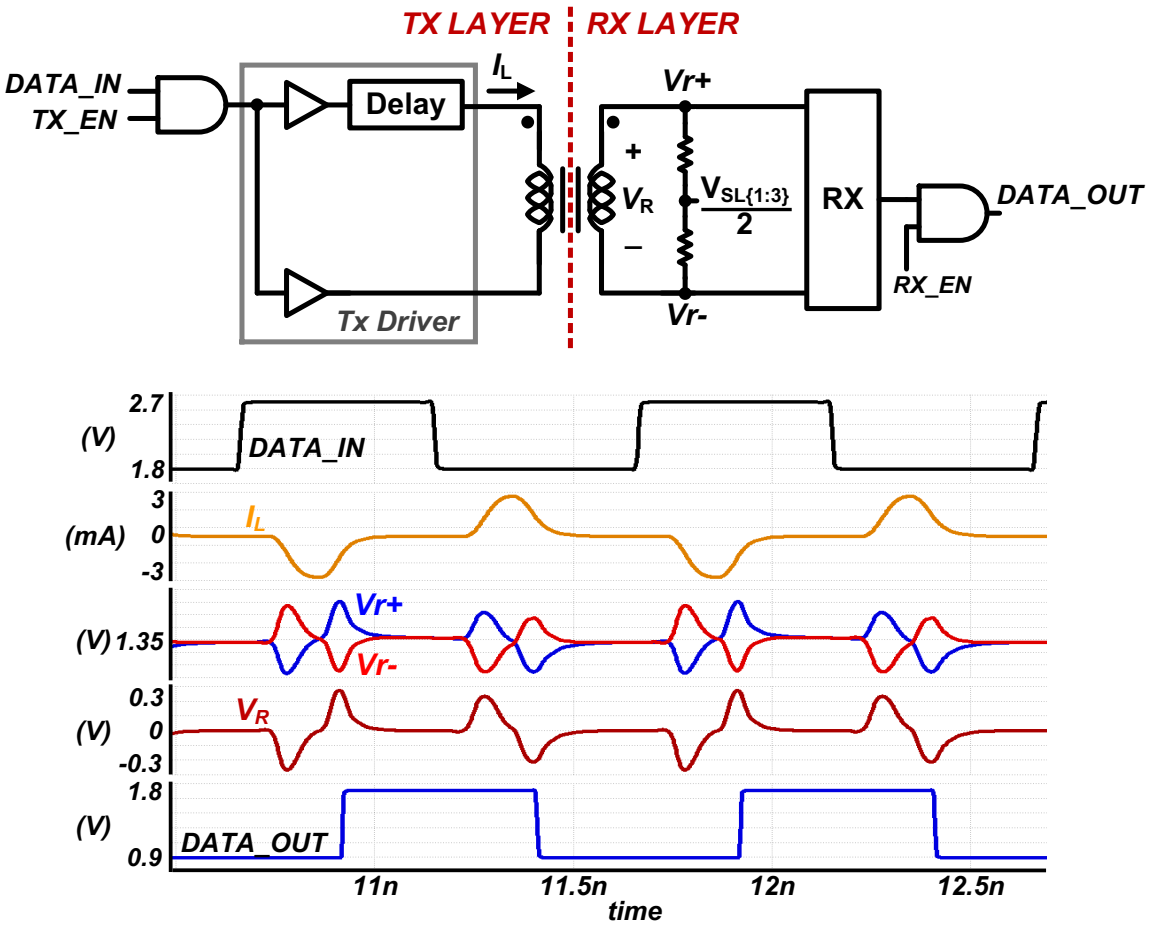


Figure 5.3: Block diagram of inductively-coupled layer shifter (top), and simulation waveforms (bottom).

capacitance (C_{C1} + C_{C2}) per wire, implemented with MoM caps on M1-M3. Tx caps occupy a total area of 24um^2 per link.

5.3 Inductively-Coupled Layer Shifter

Inductively-coupled layer shifters shown in Figure 5.3 (top), provide inherent galvanic isolation between the stacked layers. To transmit data, the Tx driver generates a current pulse on the primary coil on every data transition, creating a bi-phase mod-

Table 5.1: Summary of inductor parameters for inductively-coupled communication circuit

Metal Layers	Tx	Rx1	Rx2	Rx3
# Turns	7			
Dimension	24um			
Wire Width (pitch)	0.4um (1um)			
Inductance (L)	3.93nH	3.88nH	4.44nH	4.24nH
Resistance (R)	51Ω	193Ω	333Ω	333Ω
Mutual Inductance (M)	NA	2.83nH	2.17nH	1.77nH

ulated (BPM) voltage pulse VR across the secondary coil. The common hysteresis receiver also used in capacitive layer shifters recovers the data using the second phase of the BPM pulse. The peak and duration of VR is determined by di/dt, current pulse width, and mutual inductance [39]. The primary (Tx) coil is implemented using thick top level metals M10/M9 to minimize the series resistance of the inductor that degrades the Q factor. The secondary (Rx) coils are subsequently implemented using M8/M7, M6/M5 or M4/M3 metal layers respectively. Inductors are implemented using layout techniques to minimize the parasitic capacitance between the within-coil metal layers, as proposed in [77], and patterned ground shields are used to further improve the inductor quality [72]. While inductor alignment is an issue for inter-chip communication [42], it is not a factor for on-chip signaling, where inductors can be tightly coupled. The inductor parameters are summarized in Table 5.1.

5.4 Broadcast Communication to Multiple Layers

In addition to efficient voltage isolation, inter-layer communication circuits in voltage stacked systems must also allow for efficient broadcast of signals from one

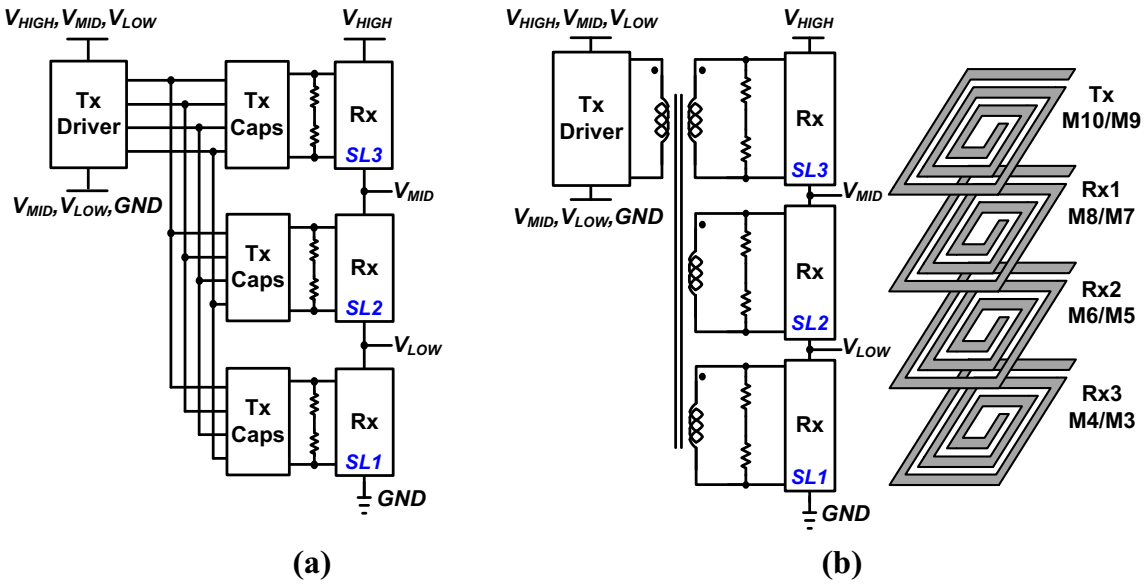


Figure 5.4: Block diagram of 1-to-3 broadcast communication for capacitively-coupled layer shifter (left) and inductively-coupled layer shifter with stacked inductor structure (right).

to multiple layers, especially to send global signals such as the clock throughout the entire chip. Although multiple 1-to-1 transmitter-receiver pairs can be used to either send signals directly to or relay the signals through multiple stack layers, these methods incur additional transmitter power, area and latency overheads. Therefore, it is desirable for a single transmitter to efficiently transmit signals directly to multiple stack layers. For the inductively-coupled layer shifter scheme presented, multiple Rx coils can couple with a single Tx coil using stacked inductor structures. Figure 5.4 presents and illustration of 1-to-3 inter-layer broadcast scheme for inductive layer shifters implemented on-chip that sends data from layer SL3 to all three stacked layers. A single Tx coil can couple to as many Rx coils as allowed by the inductor design and metal stack of the process technology. For the capacitive layer shifters, a single transmitter couples to three receivers by driving three Tx capacitor blocks.

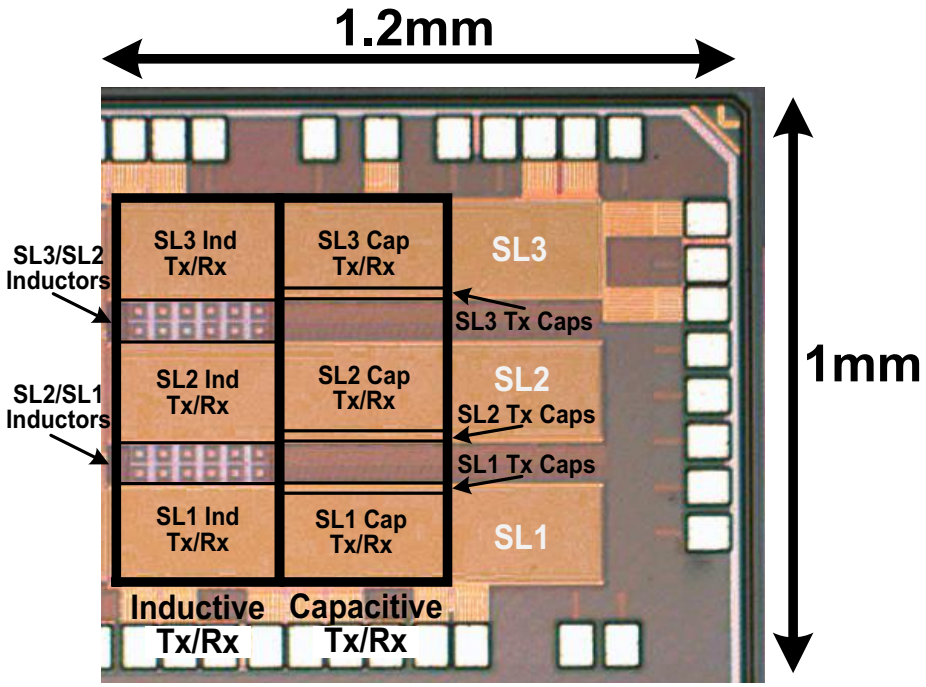
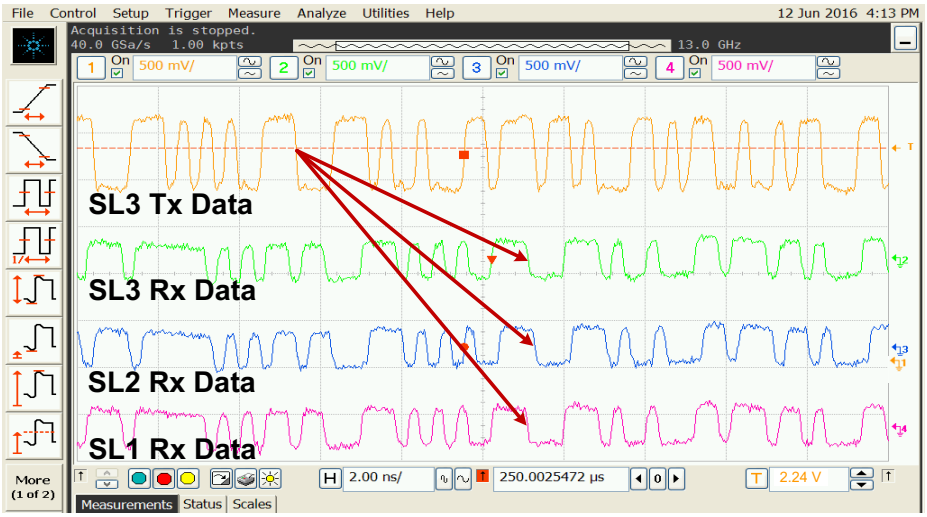


Figure 5.5: Annotated die photo of communication test chip implemented in TSMC 40G process.

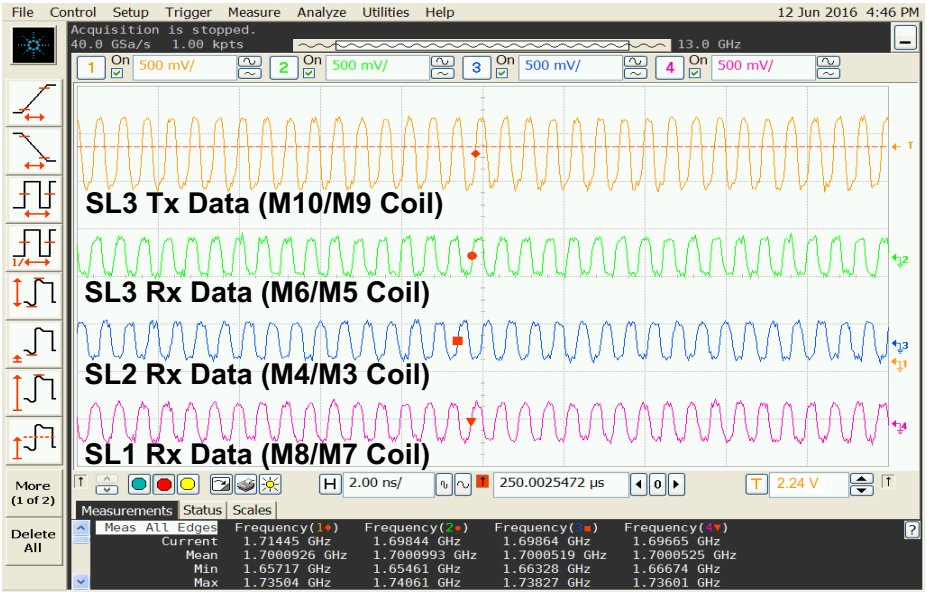
The inherent characteristics of both signaling approaches make broadcast communication possible with minimal overhead compared to full-swing signaling schemes. The inductive layer shifter can couple to multiple receivers with minimal Tx energy overhead. For the capacitive layer shifter, the combination of small capacitive load seen by the driver and low-swing pulses on the wires result in only a minor Tx energy overhead due to increased driver fanout.

5.5 Measurement Results

Figure 5.5 shows the annotated die photo of the test chip prototype implemented in TSMCs 40G process, with annotated Tx/Rx circuits as well as coupled inductors



(a)



(b)

Figure 5.6: Oscilloscope snapshots of (a) 1-to-3 capacitively-coupled broadcast communication scheme transmitting PRBS data at 3.4Gb/s at 0.9V and (b) 1-to-3 inductively-coupled broadcast communication scheme transmitting clock signals running at 1.7GHz.

and Tx capacitor banks. The coupled inductors and capacitors are placed close to the Tx layers to minimize the driver fanout and to series resistance to the inductors. To test the layer shifting communication circuits, 2^{23}-1 pseudo-random binary sequence

(PRBS) and clock signal are generated off-chip by a data-timing generator and sent directly to the transmit layer. The maximum data rate for the data timing generator is 3.4Gb/s or 1.7GHz clock frequency. Both layer shifter schemes successfully communicate PRBS data and clock signals across the stack layer boundaries at the maximum equipment data rates at 0.9V, and the maximum data rate was limited by the test equipment.

Figure 5.6 (a) presents oscilloscope waveforms of 1-to-3 layer broadcast communication for the capacitively-coupled layer shifter transmitting PRBS data at 3.4Gb/s while Figure 5.6 (b) demonstrates 1-to-3 broadcast communication for the inductively-coupled layer shifter transmitting the clock signal at 1.7GHz, with both layer shifters operating at nominal stack layer voltages of 0.9V. The waveforms demonstrate communication from layer SL3 to layers SL1, SL2, and SL3. Note that in this case, Tx circuits in layer SL3 must communicate directly to Rx circuits in SL1 demonstrating repeater-less communication between non-adjacent stack layers. As mentioned previously, all output signals are driven off-chip through a DC blocking capacitor and into the oscilloscopes 50Ω termination, which reduces the observed output voltage swing to ∼0.5V around ground. While communication between the stack layers is repeater-less, the signal must be buffered from the receiver to the output pads. The latency differences between the output signals of each layer seen in Figure 5.6 are due to the discrepancy in the distance to the output pads.

Table 5.2 presents the measured energy for both inductive and capacitive layer shifter schemes. The energy-per-bit for 1-to-1 inter-layer communication is the average energy measured over all 1-to-1 paths measured at 3.4Gb/s at 0.9V. Average

Table 5.2: Energy Comparison of Inductively- and Capacitively-coupled Communication Circuits

	Inductive		Capacitive	
	1-to-1	1-to-3	1-to-1	1-to-3
Transmitter	214 fJ/b	216 fJ/b	52 fJ/b	56 fJ/b
Receiver	71fJ/b (per layer)			
Energy/bit/link	285 fJ/b	143 fJ/b	123 fJ/b	90 fJ/b

energy for 1-to-3 inter-layer communication is also reported. In this prototype, the inductive transmitter consumes more power than its capacitive counterpart, because of the relatively large transmit current flowing through the Tx coil required to generate sufficient VR pulse. For 1-to-3 broadcast communication, increase in transmit energy is less than 5fJ/bit for both schemes compared to 1-to-1 communication, illustrating efficient multi-layer broadcast communication of the coupled schemes. Furthermore, the Tx energy of 1-to-3 broadcast communication can be amortized across the number of Rx layers, making the broadcast scheme more energy-efficient per bit per link.

The layer shifters must operate reliably with varying Tx and Rx layer voltages as internal voltages of voltage stacked systems may fluctuate due to inter-layer power mismatch [34]. Figure 5.7 presents shmoo plots obtained by testing 1-to-1 communication between layers SL1 and SL2 across a wide range of Tx and Rx layer voltages. PRBS input with constant data rate of 3.4Gb/s is transmitted throughout. At low Tx voltage, pulse-mode layer shifters have an inherent advantage over full-swing layer shifters because they only have to generate pulses sufficient to be detected by the receiver. The shmoo plot demonstrates successful capacitive communication down to 0.65V Tx voltage. The inductive scheme is less robust at lower Tx voltages, however, because generating the necessary current pulse to induce sufficient VR pulse on the

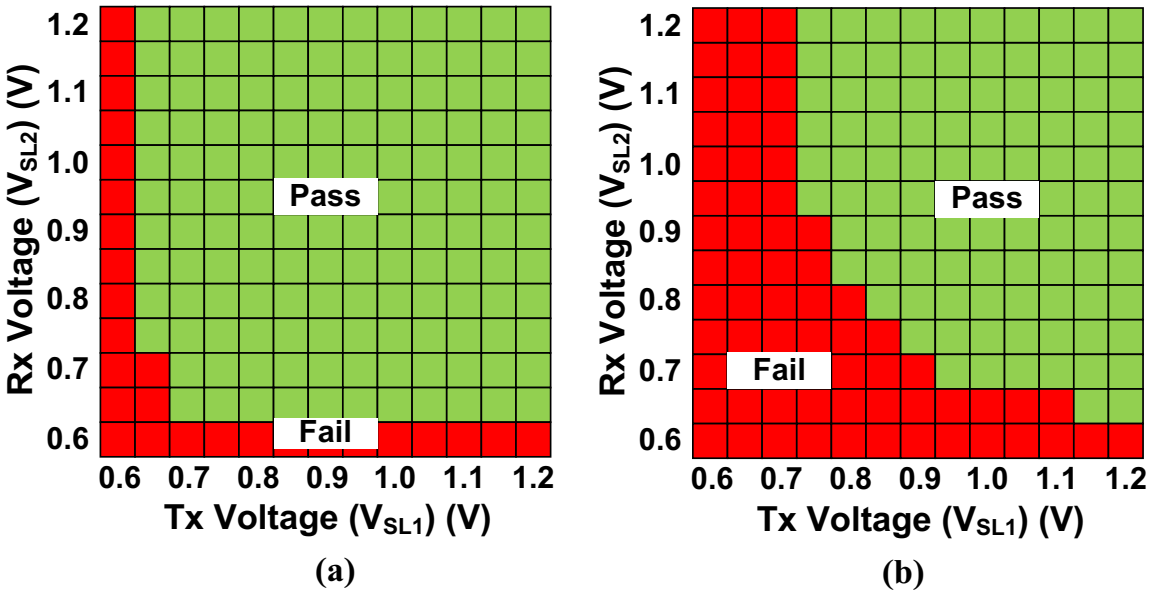


Figure 5.7: Shmoo plots for (a) capacitively-coupled and (b) inductively-coupled layer shifters with varying Tx and Rx layer voltages.

Rx coil becomes more challenging at low voltages. An ability to control the current pulse shape as proposed in [39] would improve inductive layer shifter operation at low voltages. Below 0.65V, the NMOS inputs of the hysteresis receiver cannot switch the cross-coupled PMOS pull-up transistors to produce a data transition within the pulse width.

Finally, area comparisons of the passive elements and Tx/Rx circuits, and average

Table 5.3: Area Comparison of Inductively- and Capacitively-coupled Communication Circuits

	Inductive		Capacitive	
	1-to-1	1-to-3	1-to-1	1-to-3
Tx Driver	135um^2		60um^2	
Receiver	20um^2	60um^2	20um^2	60um^2
Passives	576um^2	576um^2	24um^2	72um^2
Area/link	731um^2	257um^2	104um^2	64um^2

area per link are presented in Table 5.3. The inductive transmitter occupies a larger overall area than the capacitive transmitter due to large on-chip spiral inductor size and the larger driver size required to generate the transmit current.

Chapter 6

Conclusion and Future Directions

Voltage stacking alleviates inefficiencies related to off-chip components that hinder efficient power delivery for future computing systems, while offering the potential for highly efficient on-chip DC-DC conversion via charge recycling. However, voltage stacking suffers from internal voltage noise issues due to inter-layer activity mismatch, while adding the complexity of communication across stack layers. This thesis presented two effective voltage noise mitigation techniques, adaptive frequency clocking and a fully-integrated symmetric ladder switched-capacitor voltage regulator, and the design and implementation details of a test chip that implements a 16-core 4-way voltage-stacked system. The experimental results demonstrate effective noise mitigation, addressing the critical issue of voltage noise, while validating the high efficiency power delivery capability of voltage stacking. The attributes of voltage noise presented in this thesis demonstrates that voltage stacking will scale well with future high-throughput many-core systems, where power fluctuations of a single core will have less impact on voltage noise, while providing more opportunity for software-

level workload balancing between the cores. The effectiveness of workload balancing on overall power delivery efficiency was also verified using experimental data from the test chip.

This thesis also demonstrated efficient on-chip inter-layer signaling circuitry that leveraged coupled communication schemes previously used for chip-to-chip communication. The coupled communication scheme allows efficient communication between any stack layers and enables repeater-less long-distance signaling, which enables the application and adoption of voltage stacking in future many-core systems with high communication needs, without being hampered by large communication overheads.

Moving forward, there are various areas of research that provide ample opportunities to build upon the work presented in this thesis and further improve the performance and efficiency of voltage stacked systems:

1. **Advanced process technologies for better IVR efficiency:** For high-throughput systems that run massively parallel workloads, voltage stacking effectively hides the IVR efficiency to achieve high efficiency power delivery. However, the IVR efficiency still constitutes the lower-bound of the system-wide power delivery efficiency, and limits the efficiency for sub-optimal unbalanced workload scenarios. Better fabrication technology that can provide not only faster, more efficient transistors but high density capacitors such as deep trench capacitors demonstrated in [3, 4, 11] and/or better magnetics [17] to improve the power density and conversion efficiency of the IVR would improve the overall system-wide power delivery efficiency of voltage stacking across a wide range of workload scenarios.

2. **Circuit design techniques:** In addition to better fabrication technology, circuit advances that better regulate the voltage noise would improve the overall system efficiency. For instance, in addition to lower-bound control implemented in Section 4, additional upper-bound control can be easily implemented to keep a tighter bound on the voltage rails. Different IVR topologies, such as hybrid inductor-capacitor topologies that can overcome the regulation limitations of switched-capacitor regulators would also further improve the system efficiency.

3. **Software techniques for efficient workload balancing:** While hardware noise mitigation techniques provide the minimum voltage and reliability guarantees, and sets the lower-bound for power delivery efficiency, intelligent workload balancing to minimize the activity mismatch allows the system to approach 100% efficiency, essentially masking the hardware overheads. Using software techniques to regulate the current consumption of the system has been shown in numerous prior works to mitigate *di/dt* noise. These techniques have focused on identifying problem code segments and predicting voltage emergency events that cause sudden surge in current consumption, and using code transformation and scheduling to mitigate voltage noise [50, 51, 19, 21, 20, 26, 43, 45, 52]. Similar principles can be extended to the context of voltage stacked multicore systems, to limit the activity mismatch between stack layers.

Voltage stacking has the potential to bring significant improvements in power delivery for future high-throughput multicore systems. Moving forward, it is important to take a holistic approach to research, to pursue progress across multiple disciplines from technology to circuits and software, to fully harness the benefits of voltage

stacking as a power delivery scheme.

Bibliography

[1] International technology roadmap for semiconductors. `http://www.itrs2.net`. Accessed: 2016-08-20.

[2] Elad Alon and Mark Horowitz. Integrated regulation for energy-efficient digital circuits. *IEEE Journal of Solid-State Circuits*, 43(8):1795–1807, 2008.

[3] Toke Meyer Andersen, Florian Krismer, Johann Wal ter Kolar, Thomas Toifl, Christian Menolfi, Lukas Kull, Thomas Morf, Marcel Kossel, Matthias Brandli, Peter Buchmann, and Pier Andrea Francese. A sub-ns response on-chip switched-capacitor dc-dc voltage regulator delivering 3.7 w/mm 2 at 90% efficiency using deep-trench capacitors in 32nm soi cmos. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pages 90–91. IEEE, 2014.

[4] Toke Meyer Andersen, Florian Krismer, Johann Walter Kolar, Thomas Toifl, Christian Menolfi, Lukas Kull, Thomas Morf, Marcel Kossel, Matthias Brändii, and Pier Andrea Francese. 20.3 a feedforward controlled on-chip switched-capacitor voltage regulator delivering 10w in 32nm soi cmos. In *2015 IEEE International Solid-State Circuits Conference-(ISSCC) Digest of Technical Papers*, pages 1–3. IEEE, 2015.

[5] Ehsan K Ardestani, Rafael Trapani Possignolo, Jose Luis Briz, and Jose Renau. Managing mismatches in voltage stacking with coreunfolding. *ACM Transactions on Architecture and Code Optimization (TACO)*, 12(4):43, 2015.

[6] Shekhar Borkar. Thousand core chips: a technology perspective. In *Proceedings of the 44th annual Design Automation Conference*, pages 746–749. ACM, 2007.

[7] Shekhar Borkar and Andrew A Chien. The future of microprocessors. *Communications of the ACM*, 54(5):67–77, 2011.

[8] Thomas D Burd, Trevor Pering, Anthony J Stratakos, Robert W Brodersen, et al. A dynamic voltage scaled microprocessor system. *Solid-State Circuits, IEEE Journal of*, 35(11):1571–1580, 2000.

[9] Simone Campanoni, Kevin Brownell, Svilen Kanev, Timothy M Jones, Gu-Yeon Wei, and David Brooks. Helix-rc: an architecture-compiler co-design for automatic parallelization of irregular programs. *ACM SIGARCH Computer Architecture News*, 42(3):217–228, 2014.

[10] Gregory Andrew Chadwick. *Communication centric, multi-core, fine-grained processor architecture*. PhD thesis, Citeseer, 2013.

[11] Leland Chang et al. A fully-integrated switched-capacitor 2 1 voltage converter with regulation capability and 90% efficiency at 2.3 a/mm 2. In *2010 Symposium on VLSI Circuits*, pages 55–56, 2010.

[12] Chih Chen, HM Tong, and KN Tu. Electromigration and thermomigration in pb-free flip-chip solder joints. *Annual Review of Materials Research*, 40:531–555, 2010.

[13] Mark Edward Dean. *STRiP: a self-timed RISC processor*. PhD thesis, Stanford University, 1992.

[14] Saurabh Dighe, Sriram R Vangal, Paolo Aseron, Shasi Kumar, Tiju Jacob, Keith A Bowman, Jason Howard, James Tschanz, Vasantha Erraguntla, Nitin Borkar, et al. Within-die variation-aware dynamic-voltage-frequency-scaling with optimal core allocation and thread hopping for the 80-core teraflops processor. *IEEE Journal of Solid-State Circuits*, 46(1):184–193, 2011.

[15] Ronald G Dreslinski, Michael Wieckowski, David Blaauw, Dennis Sylvester, and Trevor Mudge. Near-threshold computing: Reclaiming moore's law through energy efficient integrated circuits. *Proceedings of the IEEE*, 98(2):253–266, 2010.

[16] Hadi Esmaeilzadeh, Emily Blem, Renee St Amant, Karthikeyan Sankaralingam, and Doug Burger. Dark silicon and the end of multicore scaling. In *Computer Architecture (ISCA), 2011 38th Annual International Symposium on*, pages 365–376. IEEE, 2011.

[17] Donald S Gardner, Gerhard Schrom, Fabrice Paillet, Tanay Karnik, and Shekhar Borkar. Integrated inductors with magnetic materials for on-chip power conversion. In *2011 IEEE Hot Chips 23 Symposium (HCS)*, pages 1–36. IEEE, 2011.

[18] Aaron Grenat, Sanjay Pant, Ravinder Rachala, and Samuel Naffziger. Adaptive clocking system for improved power efficiency in a 28nm x86-64 microprocessor. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pages 106–107. IEEE, 2014.

[19] Ed Grochowski, David Ayers, and Vivek Tiwari. Microarchitectural simulation and control of di/dt-induced power supply voltage variation. In *High-Performance Computer Architecture, 2002. Proceedings. Eighth International Symposium on*, pages 7–16. IEEE, 2002.

[20] Meeta S Gupta, Vijay Janapa Reddi, MD Smith, Gu-Yeon Wei, and DM Brooks. An event-guided approach to handling inductive noise in processors. In *Design, Automation, and Test in Europe Conference (DATE-09), Nice, France*, 2009.

[21] Meeta Sharma Gupta, Krishna K Rangan, Michael D Smith, Gu-Yeon Wei, and David Brooks. Towards a software approach to mitigate voltage emergencies. In *Low Power Electronics and Design (ISLPED), 2007 ACM/IEEE International Symposium on*, pages 123–128. IEEE, 2007.

[22] Ron Ho, Tarik Ono, Robert David Hopkins, Alex Chow, Justin Schauer, Frankie Y Liu, and Robert Drost. High speed and low energy capacitively driven on-chip wires. *IEEE Journal of Solid-State Circuits*, 43(1):52–60, 2008.

[23] Yoshifumi Ikenaga, Masahiro Nomura, Shuji Suenaga, Hideo Sonohara, Yoshitaka Horikoshi, Toshiyuki Saito, Yukio Ohdaira, Yoichiro Nishio, Tomohiro Iwashita, Miyuki Satou, et al. A 27% active-power-reduced 40-nm cmos multimedia soc with adaptive voltage scaling using distributed universal delay lines. *Solid-State Circuits, IEEE Journal of*, 47(4):832–840, 2012.

[24] Pulkit Jain, Tae-Hyoung Kim, John Keane, and Chris H Kim. A multi-story power delivery technique for 3d integrated circuits. In *Proceedings of the 2008 international symposium on Low Power Electronics & Design*, pages 57–62. ACM, 2008.

[25] Rinkle Jain, Bibiche M Geuskens, Stephen T Kim, Muhammad M Khellah, Jitendra Kulkarni, James W Tschanz, and Vivek De. A 0.45–1 v fully-integrated distributed switched capacitor dc-dc converter with high density mim capacitor in 22 nm tri-gate cmos. *Solid-State Circuits, IEEE Journal of*, 49(4):917–927, 2014.

[26] Russ Joseph, David Brooks, and Margaret Martonosi. Control techniques to eliminate voltage emergencies in high performance processors. In *High-Performance Computer Architecture, 2003. HPCA-9 2003. Proceedings. The Ninth International Symposium on*, pages 79–90. IEEE, 2003.

[27] Ulya R Karpuzcu, Abhishek Sinkar, Nam Sung Kim, and Josep Torrellas. Energysmart: toward energy-efficient manycores for near-threshold computing. In *High Performance Computer Architecture (HPCA2013), 2013 IEEE 19th International Symposium on*, pages 542–553. IEEE, 2013.

[28] Kapil Kesarwani, Christopher Schaef, Charles R Sullivan, and Jason T Stauth. A multi-level ladder converter supporting vertically-stacked digital voltage domains. In *Applied Power Electronics Conference and Exposition (APEC), 2013 Twenty-Eighth Annual IEEE*, pages 429–434. IEEE, 2013.

[29] Wonyoung Kim, David Brooks, and Gu-Yeon Wei. A fully-integrated 3-level dc-dc converter for nanosecond-scale dvfs. *Solid-State Circuits, IEEE Journal of*, 47(1):206–219, 2012.

[30] Wonyoung Kim, Meeta S Gupta, Gu-Yeon Wei, and David Brooks. System level analysis of fast, per-core dvfs using on-chip switching regulators. In *2008 IEEE 14th International Symposium on High Performance Computer Architecture*, pages 123–134. IEEE, 2008.

[31] George Kurian, Jason E Miller, James Psota, Jonathan Eastep, Jifeng Liu, Jurgen Michel, Lionel C Kimerling, and Anant Agarwal. Atac: a 1000-core cache-coherent processor with on-chip optical network. In *Proceedings of the 19th international conference on Parallel architectures and compilation techniques*, pages 477–488. ACM, 2010.

[32] Hanh-Phuc Le, Seth R Sanders, and Elad Alon. Design techniques for fully integrated switched-capacitor dc-dc converters. *Solid-State Circuits, IEEE Journal of*, 46(9):2120–2131, 2011.

[33] Sae Kyu Lee, David Brooks, and Gu-Yeon Wei. Evaluation of voltage stacking for near-threshold multicore computing. In *Proceedings of the 2012 ACM/IEEE international symposium on Low power electronics and design*, pages 373–378. ACM, 2012.

[34] Sae Kyu Lee, Tao Tong, Xuan Zhang, David Brooks, and Gu-Yeon Wei. A 16-core voltage-stacked system with an integrated switched-capacitor dc-dc converter. In *VLSI Circuits (VLSIC), 2015 Symposium on*, pages C318–C319. IEEE, 2015.

[35] Charles R Lefurgy, Alan J Drake, Michael S Floyd, Malcolm S Allen-Ware, Bishop Brock, Jose Tierno, John B Carter, Robert W Berry, et al. Active guardband management in power7+ to save energy and maintain reliability. *Micro, IEEE*, 33(4):35–45, 2013.

[36] Yu-Shiang Lin, Dennis Sylvester, and David Blaauw. Alignment-independent chip-to-chip communication for sensor applications using passive capacitive signaling. *IEEE Journal of Solid-State Circuits*, 44(4):1156–1166, 2009.

[37] Kenneth W Mai, Toshihiko Mori, Bharadwaj S Amrutur, Ron Ho, Bennett Wilburn, Mark A Horowitz, Isao Fukushi, Tetsuo Izawa, and Shin Mitarai. Low-power sram design using half-swing pulse-mode techniques. *IEEE Journal of Solid-State Circuits*, 33(11):1659–1671, 1998.

[38] Josiah McClurg, Yujia Zhang, Jeff Wheeler, and Robert Pilawa-Podgurski. Rethinking data center power delivery: Regulating series-connected voltage domains in software. In *Power and Energy Conference at Illinois (PECI), 2013 IEEE*, pages 147–154. IEEE, 2013.

[39] Noriyuki Miura, Hiroki Ishikuro, Takayasu Sakurai, and Tadahiro Kuroda. A 0.14 pj/b inductive-coupling inter-chip data transceiver with digitally-controlled precise pulse shaping. In *54th IEEE International Solid-State Circuits Conference, ISSCC 2007*, 2007.

[40] Noriyuki Miura, Daisuke Mizoguchi, Mari Inoue, Kiichi Niitsu, Yoshihiro Nakagawa, Masamoto Tago, Muneo Fukaishi, Takayasu Sakurai, and Tadahiro Kuroda. A 1 tb/s 3 w inductive-coupling transceiver for 3d-stacked inter-chip clock and data link. *IEEE Journal of Solid-State Circuits*, 42(1):111–122, 2007.

[41] Makoto Nagata, Takeshi Okumoto, and Kazuo Taki. A built-in technique for probing power supply and ground noise distribution within large-scale digital integrated circuits. *IEEE Journal of Solid-State Circuits*, 40(4):813–819, 2005.

[42] Kiichi Niitsu, Yoshinori Kohama, Yasufumi Sugimori, Kazutaka Kasuga, Kenichi Osada, Naohiko Irie, Hiroki Ishikuro, and Tadahiro Kuroda. Modeling and experimental verification of misalignment tolerance in inductive-coupling inter-chip link for low-power 3-d system integration. *IEEE transactions on very large scale integration (VLSI) systems*, 18(8):1238–1243, 2010.

[43] Mondira Deb Pant, Pankaj Pant, D Scott Wills, and Vivek Tiwari. An architectural solution for the inductive noise problem due to clock-gating. In *Low Power Electronics and Design, 1999. Proceedings. 1999 International Symposium on*, pages 255–257. IEEE, 1999.

[44] Nathaniel Pinckney, Korey Sewell, Ronald G Dreslinski, David Fick, Trevor Mudge, Dennis Sylvester, and David Blaauw. Assessing the performance limits of parallelized near-threshold computing. In *Proceedings of the 49th Annual Design Automation Conference*, pages 1147–1152. ACM, 2012.

[45] Michael D Powell and TN Vijaykumar. Pipeline muffling and a priori current ramping: architectural techniques to reduce high-frequency inductive noise. In *Low Power Electronics and Design, 2003. ISLPED'03. Proceedings of the 2003 International Symposium on*, pages 223–228. IEEE, 2003.

[46] Saravanan Rajapandian, Kenneth L Shepard, Peter Hazucha, and Tanay Karnik. High-voltage power delivery through charge recycling. *Solid-State Circuits, IEEE Journal of*, 41(6):1400–1410, 2006.

[47] Saravanan Rajapandian, Zheng Xu, and Kenneth L Shepard. Implicit dc-dc downconversion through charge-recycling. *Solid-State Circuits, IEEE Journal of*, 40(4):846–852, 2005.

[48] Sabela Ramos and Torsten Hoefler. Modeling communication in cache-coherent smp systems: a case-study with xeon phi. In *Proceedings of the 22nd international symposium on High-performance parallel and distributed computing*, pages 97–108. ACM, 2013.

[49] Brandon Reagen, Robert Adolf, Yakun Sophia Shao, Gu-Yeon Wei, and David Brooks. Machsuite: Benchmarks for accelerator design and customized architectures. In *Workload Characterization (IISWC), 2014 IEEE International Symposium on*, pages 110–119. IEEE, 2014.

[50] Vijay Janapa Reddi, Meeta S Gupta, Glenn Holloway, Gu-Yeon Wei, Michael D Smith, and David Brooks. Voltage emergency prediction: Using signatures to reduce operating margins. In *2009 IEEE 15th International Symposium on High Performance Computer Architecture*, pages 18–29. IEEE, 2009.

[51] Vijay Janapa Reddi, Svilen Kanev, Wonyoung Kim, Simone Campanoni, Michael D Smith, Gu-Yeon Wei, and David Brooks. Voltage smoothing: Characterizing and mitigating voltage noise in production processors via software-guided thread scheduling. In *2010 43rd Annual IEEE/ACM International Symposium on Microarchitecture*, pages 77–88. IEEE, 2010.

[52] Vijay Janapa Reddi, Svilen Kanev, Wonyoung Kim, Simone Campanoni, Michael D Smith, Gu-Yeon Wei, and David M Brooks. Voltage noise in production processors. *IEEE micro*, 31(1):20–28, 2011.

[53] Mitsuko Saito, Yasufumi Sugimori, Yoshinori Kohama, Yoichi Yoshida, Noriyuki Miura, Hiroki Ishikuro, Takayasu Sakurai, and Tadahiro Kuroda. 2 gb/s 15 pj/b/chip inductive-coupling programmable bus for nand flash memory stacking. *IEEE Journal of Solid-State Circuits*, 45(1):134–141, 2010.

[54] Seth R Sanders, Elad Alon, Hanh-Phuc Le, Michael D Seeman, Mervin John, and Vincent W Ng. The road to fully integrated dc–dc conversion via the switched-capacitor approach. *IEEE Transactions on Power Electronics*, 28(9):4146–4155, 2013.

[55] Steven L Scott. Synchronization and communication in the t3e multiprocessor. In *ACM SIGPLAN Notices*, volume 31, pages 26–36. ACM, 1996.

[56] Jae-sun Seo, Ron Ho, Jon Lexau, Michael Dayringer, Dennis Sylvester, and David Blaauw. High-bandwidth and low-energy on-chip signaling with adaptive pre-emphasis in 90nm cmos. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, pages 182–183. IEEE, 2010.

[57] Pradeep S Shenoy and Philip T Krein. Differential power processing for dc systems. *Power Electronics, IEEE Transactions on*, 28(4):1795–1806, 2013.

[58] J Gregory Steffan, Christopher B Colohan, Antonia Zhai, and Todd C Mowry. Improving value communication for thread-level speculation. In *High-Performance Computer Architecture, 2002. Proceedings. Eighth International Symposium on*, pages 65–75. IEEE, 2002.

[59] Michael Bedford Taylor, Jason Kim, Jason Miller, David Wentzlaff, Fae Ghodrat, Ben Greenwald, Henry Hoffman, Paul Johnson, Jae-Wook Lee, Walter Lee, et al. The raw microprocessor: A computational fabric for software circuits and general-purpose programs. *IEEE micro*, 22(2):25–35, 2002.

[60] Carlos Tokunaga, Joseph F Ryan, Charles Augustine, Jaydeep P Kulkarni, Yi-Chun Shih, Stephen T Kim, Rinkle Jain, Keith Bowman, Arijit Raychowdhury, Muhammad M Khellah, et al. A graphics execution core in 22nm cmos featuring adaptive clocking, selective boosting and state-retentive sleep. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pages 108–109. IEEE, 2014.

[61] Tao Tong, Sae Kyu Lee, Xuan Zhang, David Brooks, and Gu-Yeon Wei. A fully integrated reconfigurable switched-capacitor dc-dc converter with four stacked output channels for voltage stacking applications. *Solid-State Circuits, IEEE Journal of*, 51(9), 2016.

[62] Tao Tong, Xuan Zhang, Wonyoung Kim, David Brooks, and Gu-Yeon Wei. A fully integrated battery-connected switched-capacitor 4:1 voltage regulator with 70% peak efficiency using bottom-plate charge recycling. In *Custom Integrated Circuits Conference (CICC), 2013 IEEE*, pages 1–4. IEEE, 2013.

[63] Kazunori Ueda, Fukashi Morishita, Shunsuke Okura, Leona Okamura, Tatsuhiko Yoshihara, and Keisuke Arimoto. Low-power on-chip charge-recycling dc-dc conversion circuit and system. *Solid-State Circuits, IEEE Journal of*, 48(11):2608–2617, 2013.

[64] András Vajda. *Programming many-core chips*. Springer Science & Business Media, 2011.

[65] Rob F Van der Wijngaart, Timothy G Mattson, and Werner Haas. Light-weight communications on intel's single-chip cloud computer processor. *ACM SIGOPS Operating Systems Review*, 45(1):73–83, 2011.

[66] Gu-Yeon Wei and Mark Horowitz. A fully digital, energy-efficient, adaptive power-supply regulator. *Solid-State Circuits, IEEE Journal of*, 34(4):520–528, 1999.

[67] JD Wu, PJ Zheng, CW Lee, SC Hung, and JJ Lee. A study in flip-chip ubm/bump reliability with effects of snpb solder composition. *Microelectronics reliability*, 46(1):41–52, 2006.

[68] Qiang Wu, Philo Juang, Margaret Martonosi, and Douglas W Clark. Voltage and frequency control with adaptive reaction time in multiple-clock-domain processors. In *11th International Symposium on High-Performance Computer Architecture*, pages 178–189. IEEE, 2005.

[69] Hiroyuki Yamauchi, Hironori Akamatsu, and Tsutomu Fujita. An asymptotically zero power charge-recycling bus architecture for battery-operated ultrahigh data rate ulsi's. *IEICE transactions on electronics*, 78(6):671–679, 1995.

[70] Byung-Do Yang and Lee-Sup Kim. A low-power rom using charge recycling and charge sharing techniques. *IEEE Journal of Solid-State Circuits*, 38(4):641–653, 2003.

[71] Yoichi Yoshida, Noriyuki Miura, and Tadahiro Kuroda. A 2 gb/s bi-directional inter-chip data transceiver with differential inductors for high density inductive channel array. *IEEE Journal of Solid-State Circuits*, 43(11):2363–2369, 2008.

[72] C Patrick Yue and S Simon Wong. On-chip spiral inductors with patterned ground shields for si-based rf ics. *IEEE Journal of solid-state circuits*, 33(5):743–752, 1998.

[73] Runjie Zhang, Kaushik Mazumdar, Brett H Meyer, Ke Wang, Kevin Skadron, and Mircea Stan. A cross-layer design exploration of charge-recycled power-delivery in many-layer 3d-ic. In *Proceedings of the 52nd Annual Design Automation Conference*, page 133. ACM, 2015.

[74] Runjie Zhang, Kaushik Mazumdar, Brett H Meyer, Ke Wang, Kevin Skadron, and Mircea R Stan. Transient voltage noise in charge-recycled power delivery networks for many-layer 3d-ic. In *Low Power Electronics and Design (ISLPED), 2015 IEEE/ACM International Symposium on*, pages 152–158. IEEE, 2015.

[75] Xuan Zhang, Tao Tong, David Brooks, and Gu-Yeon Wei. Supply-noise resilient adaptive clocking for battery-powered aerial microrobotic system-on-chip in 40nm cmos. In *Custom Integrated Circuits Conference (CICC), 2013 IEEE*, pages 1–4. IEEE, 2013.

[76] Brian Zimmer, Yunsup Lee, Alberto Puggelli, Jaehwa Kwak, Ruzica Jevtic, Ben Keller, Stevo Bailey, Milovan Blagojevic, Pi-Feng Chiu, Hanh-Phuc Le, et al. A risc-v vector processor with tightly-integrated switched-capacitor dc-dc converters in 28nm fdsoi. In *VLSI Circuits (VLSI Circuits), 2015 Symposium on*, pages C316–C317. IEEE, 2015.

[77] Alireza Zolfaghari, Andrew Chan, and Behzad Razavi. Stacked inductors and transformers in cmos technology. *IEEE Journal of Solid-State Circuits*, 36(4):620–628, 2001.