A 16-core voltage-stacked system with an integrated switched-capacitor DC-DC converter

Sae Kyu Lee*, Tao Tong*, Xuan Zhang, David Brooks, Gu-Yeon Wei
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, U.S.A

Abstract

A 16-core voltage-stacked IC integrated with a switched-capacitor DC-DC converter demonstrates efficient power delivery. To overcome inter-layer voltage noise issues, the test chip implements and evaluates the benefits of self-timed clocking and clock-phase interleaving. The integrated converter offers minimum voltage guarantees and further reduces voltage noise.

Introduction

Voltage stacking is an alternative power delivery solution for multicore processors that stacks voltage domains in series and delivers a single high voltage to the cores rather than delivering current in parallel at a low voltage [1][2]. Recycling charge through stack layers reduces overall current draw of the chip, alleviates power delivery impedance requirements, and obviates high step-down voltage regulators. However, power consumption mismatch between stack layers translates to inter-layer voltage noise that can degrade performance and energy efficiency. This paper presents a 16-core 4-way stacked test chip with self-timed clocking, clock-phase interleaving, and *a fully-integrated, symmetric-ladder, switched-capacitor converter* (IVR) to reduce inter-layer voltage noise and improve throughput. Measured results show these techniques offer power delivery benefits when voltage stacking multiple cores for high-throughput systems.

Test Chip Architecture

Fig. 1 presents an overview of the test chip implemented in TSMC's 40G process, which uses triple-wells to isolate cores in each stack layer from body-bias effects. The 4x4 stacked array and IVR operate off of a single 3.6V V_{IN}. Each *core* contains a 32-bit Intel Siskiyou Peak processor—a 5-stage, single-issue, pipeline that implements a subset of the IA instruction set—and 4KB SRAM for instructions and data. An external clock source can drive all cores for fully synchronous fixed-frequency operation (FFClk) or cores in each layer share digitally-configured ring oscillators (DCRO), shown in Fig. 2, which track worst-case logic delay paths with respect to local supply voltage fluctuations, thereby enabling per-layer, self-timed clocking (STClk) to improve throughput. Each core further includes clock-phase interleaving (ClkInt) that provides 90° out-of-phase clock signals to cores in each layer to smooth out high-frequency voltage noise. For communication between stack layers, initializable layer shifters couple signals across MoM capacitors.

Ideally, when power consumption in all layers perfectly matches, voltage stacking evenly distributes internal voltage rails to 0.9V per layer. While differing core activities and corresponding power imbalances lead to voltage noise, the companion IVR absorbs inter-layer power imbalances in the 4x4 core array and regulates the internal voltage rails (V_{UPP}, V_{MID}, and V_{LOW}). The symmetric ladder IVR topology is a natural choice for voltage stacking, because—in a sense—it too employs voltage stacking. When combined with voltage-stacked cores, the IVR operates as a multi-output converter that simultaneously regulates all four layer voltages, but must only compensate power differences between stack layers, resulting in smaller size and incurring IVR losses only for a portion of the total power consumed by the cores.

To reduce voltage ripple, the IVR comprises ten clock-interleaved SC ladder units (Fig. 1), driven by two non-overlapping clocks Φ1 and Φ2, to pull current from V_{IN} and redistribute charge across the four layers [3]. The stacked topology enables using thin-oxide transistors for its power switches and flying capacitors to improve density and conversion efficiency. The IVR contains 4.5nF of total flying capacitance, 12C per SC ladder unit, and 1C = 37.5pF. Fig. 3 illustrates the IVR's lower-bound control loop used to regulate the layer voltages ($V_{L\{4:1\}}$). Four 2.5GHz clocked voltage comparators detect whether any layer voltage falls below a minimum level set by voltage references locally generated off of each layer's lower rail. The feedback control logic combines comparator outputs and produces switching requests for the ten SC ladder units, each switching at a maximum of 250MHz, to supply and redistribute charge across the layers to balance out stacked voltage levels.

The optimum allocation of flying capacitor resources depends on which layer pulls more current from the outputs. The two different capacitor allocations in Fig. 4 (upper left) each maximize conversion efficiency for their respective load current scenarios. Consequently, each SC ladder unit includes twelve switchable capacitor banks (*cap-bank*) for configurability from 1C to 2.5C per layer. A pair of thin-oxide flying inverters implements the reconfigurable switches SW1 and SW2. Since this reconfigurability targets slow system-wide load imbalances, switching losses to drive SW1 and SW2 are small. Fig. 4 (lower left) plots measured conversion efficiencies of the IVR with respect to power delivered to a single layer set via on-chip current loads. The IVR achieves higher efficiency and supports higher output power for middle layers L2 and L3, consistent with inter-layer charge flow analysis. Moreover, imposing equal flying capacitor sizes, L1 (Reconf. off), degrades conversion efficiency and reduces maximum deliverable output power.

Measurement Results

To evaluate benefits of added features that mitigate throughput and voltage noise issues associated with voltage stacking, we measured power and frequency of the voltage-stacked cores running different combinations of 16 kernels from a suite that offer diverse workload behaviors (matrix-multiply, Viterbi, sort, etc.). Measured per-core dynamic power ranges from 7.519 mW (KMP) to 8.924mW (KNN) when run on a single core operating at 0.9V and 250MHz. Each layer's DCRO and clock distribution logic consumes 4.041mW at 0.9V and 250MHz. When completely idle, the test chip draws ~10.3mA of leakage current from the 3.6V V_{IN}. On-chip voltage monitoring circuits provide visibility to internal voltage rails.

To first understand inherent attributes of voltage stacking, i.e., STClk, ClkInt, and IVR all off, Fig. 5 (left) compares snapshots of measured V_{L1} for two workload scenarios running FFClk. In the balanced case, all cores execute KMP (a lower-power control-intensive string matching kernel) and all layer voltages evenly subdivide to 0.9V. In the unbalanced case, only the cores in L1 execute KNN (a bursty higher-power compute-intensive molecular dynamics kernel) while all other cores execute KMP. This leads to pronounced voltage droops in V_{L1} as equal current must flow through the layers. The other plots in Fig. 5 show cumulative improvements in voltage noise by successively enabling STClk, ClkInt, and IVR, for the same unbalanced scenario. IVR V_{REF} is set to 850mV for all measurements, unless specified otherwise. Histogram plots (in Fig. 6) of V_{L1} collected over a 1ms window further quantifies the improvements in worst-case voltage from 764mV to 796mV with ClkInt on, and then to 822mV with IVR on.

Worst-case voltage droops are especially problematic for fixed-frequency operation, because they require large timing margins, limiting F_{MAX} to <165MHz for V_{MIN}≈734mV. In contrast, self-timed clocking requires much smaller timing margins by tracking voltage fluctuations. This improves overall performance, confirmed by the boxplots (a distribution's average, 50%, min, and max) of L1 DCRO frequency measurements (over 1ms) in Fig. 7. For a system with STClk on, normalized throughput improves by 50% and normalized energy-delay product (EDP) reduces by 40 percentage points. Turning on ClkInt and IVR further improves system-level throughput, but EDP degrades slightly due to ClkInt logic power and IVR losses. For completeness, Fig. 8 presents boxplots of the four layer voltages, measured over a 1ms window, for the unbalanced scenario run under three test conditions discussed above. While self-timed clocking provides most of the throughput improvements, noise reduction via the IVR provides important guarantees, e.g., meet minimum latency requirements and/or guarantee minimum voltage requirements for SRAM stability, which is 0.7V in this design.

Turning on the IVR of course comes with a power penalty. To quantify its losses, we ran 300 different program scenarios, where the

* These Authors have equal contributions

four cores in each layer run one of four kernels randomly chosen from the pool of 16. Fig. 9 provides a histogram of *system-wide power delivery efficiency*, defined and calculated as the ratio of the aggregate average power consumed by all 16 cores (based on pre-characterized, per-program, core power measurements) over the total power delivered to the test chip via V_{IN} at 3.6V. For 16 balanced scenarios executing a single benchmark in all cores (highlighted with a different color), efficiency is greater than 99%, because the IVR rarely turns on. For all other scenarios, efficiency is still high (>97%) because IVR losses only apply to differential power between the layers. Finally, Fig. 10 compares measured layer voltages for a larger load transient event with and without the IVR on (STClk and ClkInt are always on). All cores initially run the KMP kernel. Then, two cores in each of the three layers, L1, L2, and L4, turn off via clock gating. With the IVR off, V_{L3} drops due to its higher activity and all other layer voltages increase for equal current to flow through all layers. In contrast, the IVR regulates all layer voltages with respect to V_{REF} = 850mV and supplies extra power for the two additional cores in L3 and slightly higher power in L1 due to an elevated voltage. Even for this grossly unbalanced condition, system-wide power delivery efficiency is 96.3%, because the IVR only supplies a small portion of the test chip's total power. Fig. 11 presents the die photo of the test chip.

Acknowledgements

The authors thank TSMC's University Shuttle Program for chip fabrication and Intel Corporation for Siskiyou Peak IP. This work was supported in part by NSF (CCF-0903437 and CCF-1218298) and DARPA (HR0011-13-C-0022).

References

[1] S. Rajapandian *et al.*, IEEE JSSC, 2006, vol. 31, pp. 1400-1410

[2] S. Lee *et al.*, ISLPED, 2012, pp. 373-378

[3] M. Seeman *et al.*, IEEE TPEL, 2008, vol. 23, pp. 841-851

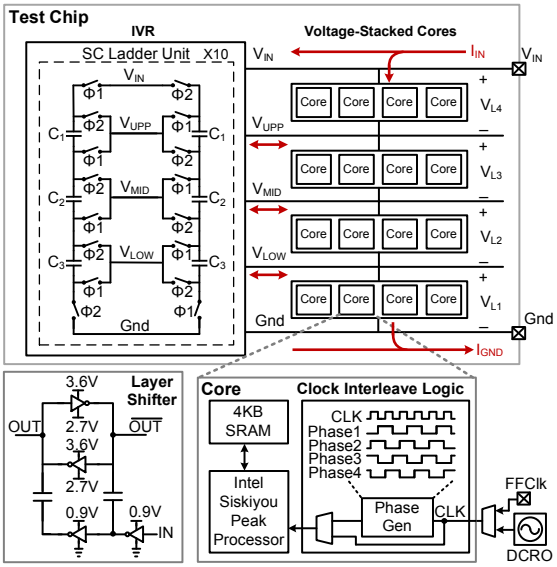

Fig. 1: Block diagram of voltage-stacked system diagram showing the 16 4-way stacked cores and IVR.

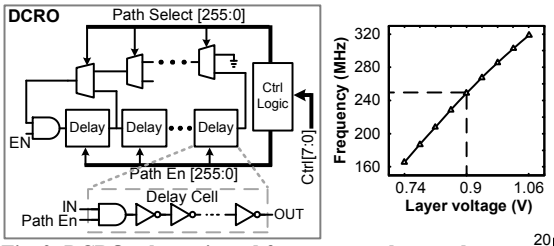

Fig. 2: DCRO schematic and frequency vs. layer voltage.

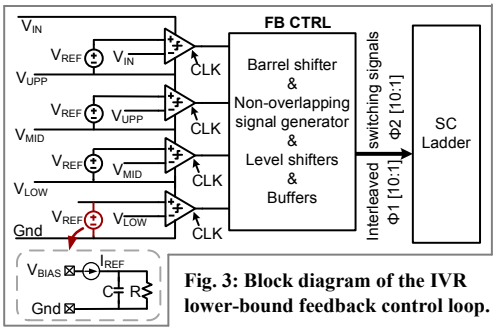

Fig. 3: Block diagram of the IVR lower-bound feedback control loop.

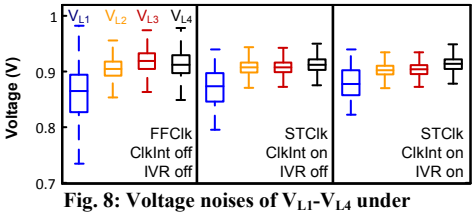

Fig. 8: Voltage noises of V_{L1}-V_{L4} under different testing conditions.

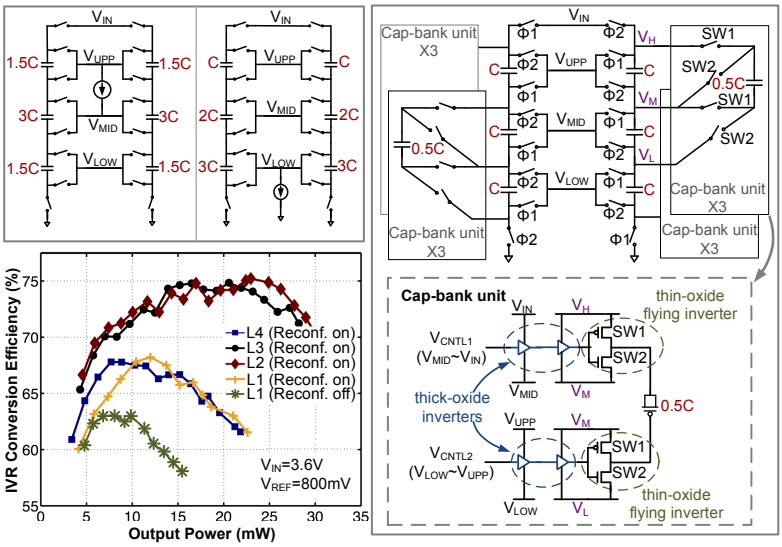

Fig. 4: Examples of optimized flying capacitor allocations for different load conditions (upper left), implementation of the reconfigurable SC ladder (right), and measured IVR efficiencies when only one layer consumes current (lower left).

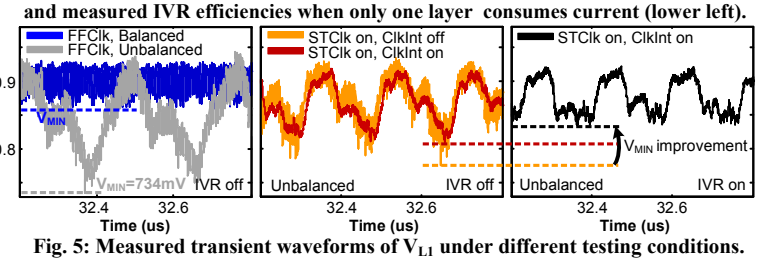

Fig. 5: Measured transient waveforms of V_{L1} under different testing conditions.

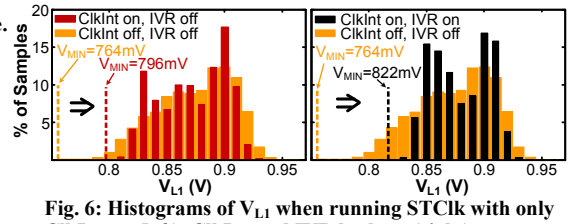

Fig. 6: Histograms of V_{L1} when running STClk with only ClkInt on (left), ClkInt and IVR both on (right).

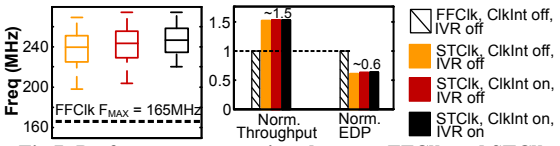

Fig.7: Performance comparison between FFClk and STClk.

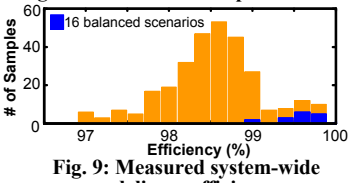

Fig. 9: Measured system-wide power delivery efficiency.

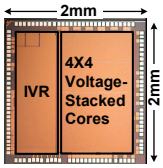

Fig. 11: Die Photo.

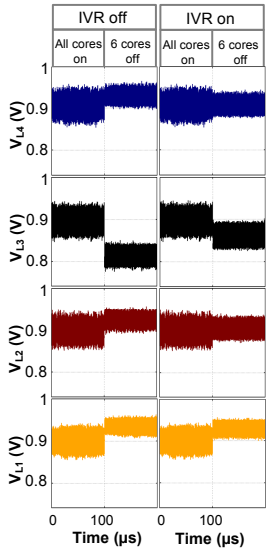

Fig. 10: Load transient waveforms via clock gating.