A single-VDD half-clock-tolerant fine-grained dynamic voltage scaling pipeline

Zhou, Rong; Chong, Kwen-Siong; Lin, Tong; Gwee, Bah Hwee; Chang, Joseph Sylvester

2015

Zhou, R., Chong, K.-S., Lin, T., Gwee, B. H., & Chang, J. S. (2015). A single-VDD half-clock-tolerant fine-grained dynamic voltage scaling pipeline. 2015 IEEE International Symposium on Circuits and Systems (ISCAS), 2589-2592.

<https://hdl.handle.net/10356/82870>

<https://doi.org/10.1109/ISCAS.2015.7169215>

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [<http://dx.doi.org/10.1109/ISCAS.2015.7169215>].

Downloaded on 26 Apr 2023 20:41:05 SGT

A Single- V_{DD} Half-Clock-Tolerant Fine-Grained Dynamic Voltage Scaling Pipeline

Rong Zhou*, Kwen-Siong Chong, Tong Lin, Bah-Hwee Gwee, and Joseph S. Chang
Nanyang Technological University, Singapore
*ZHOU0173@e.ntu.edu.sg

Abstract—We propose a novel dynamic voltage scaling (DVS) pipeline with three significant attributes. First, it features a fine-grained DVS which innately attempts to power most of the circuits therein at low voltages, and when the speed is beneath the requirement, to scale up the voltage. Second, it supports fast-transition DVS within one-and-a-half clock duration per operation, and its operation remains error-free during that duration; we define such attribute as half-clock-tolerant. Third, it consists of a single power source (single- V_{DD}) which supports three voltage scales (1.2V, 0.8V and 0.5V) for power/speed trade-offs, and has standardized 1.2V output to seamlessly interface with other proposed/conventional pipelines. These attributes are achieved due to the embodiment of a DVS power unit, asynchronous building blocks to control/synchronize the operation, a dual-rail critical path to innately detect the completion of the operation, and level shifters to standardize the output voltage. We demonstrate our proposed pipeline by designing a multiplier embodied in a Fast Fourier Transform processor (@65nm CMOS). We show that the multiplier based on our proposed pipeline, on average, is 1.94× more power-efficient than that based on a conventional pipeline.

Keywords— *asynchronous; critical path; dual-rail; dynamic voltage scaling; fast transition, fine-grained; timing-tolerant*

I. INTRODUCTION

Dynamic Voltage Scaling (DVS) is one of the most widely used techniques to trade-off speed for power, or vice-versa [1]-[3]. With DVS, the power dissipation can be reduced by scaling down the supply voltage V_{DD} when the operating workload is low. The converse is true when the operating workload is high. Generally, the degree of DVS can be done in an either course-grained or fine-grained manner. The coarse-grained DVS can be easily achieved by using a dedicated DC-DC converter to adjust V_{DD} of a cluster of blocks/IPs. The limitations of the course-grained DVS are limited power efficiency and slow V_{DD} transition (e.g. 10s-100s μ s) [4]. On the other hand, the fine-grained DVS can be achieved by using integrated power switches applied to a smaller block enclosed by an either multi- V_{DD} or single- V_{DD} power source. Between the multi- V_{DD} and single- V_{DD} DVS, the former could potentially have better power efficiency (e.g. via voltage dithering [5]) at the cost of higher routing overhead. The latter would be more advantageous in terms of simple implementation, fast-transition [4], and small-area. In this study, we focus on the single- V_{DD} fine-grained DVS.

A critical issue for DVS is to scale V_{DD} according to the

corresponding workload and yet to ensure the correct operation. The challenge of ensuring the correct operation arises from timing variations due to voltage scaling and to some extent to process and temperature variations. The reported techniques to mitigate erroneous operation during DVS include self-timed asynchronous operation [6], the error detection/recovery techniques such as Razor [7], the careful parameter pre-characterization (e.g. the frequency-voltage profile is stored in a lookup-table on-chip), throughput monitoring, critical path replicas [8]-[10], and the simplistic method by inserting large timing margins to cater for the timing uncertainty. Although the self-timed asynchronous operation and the error detection/recovery techniques appear to be robust, they require somewhat large overheads and the latter might have delay penalty (especially when the error detection rate is high). For the other techniques, their power-efficiency and robustness are strongly dependent on the specific controller implementation (to transit from one voltage level to another), and most of them are more appropriate to be realized for the coarse-grained DVS (instead of the fine-grained DVS) largely due to the inherent timing uncertainty.

In this paper, we propose a single- V_{DD} robust (timing-tolerant) fine-grained DVS pipeline for low power low-to-mid speed applications. The proposed pipeline features the following attributes. First, it will innately scale down the virtual supply voltage $V-V_{DD}$ (that powers most of the circuits therein) to the possible lowest voltage for low power dissipation, and to scale up $V-V_{DD}$ when the speed is beneath the requirement. Second, it supports fine-grained DVS within one clock (per operation) and is tolerant up to 1½ clock when its computing speed is beneath than a clock speed. The tolerance of the additional ½ clock is to prevent any possible erroneous operation and to attempt to recover the speed within the ½ clock. Third, it is powered by a single- V_{DD} , and its voltage scaling achieved by embedded three power switch paths (for 1.2V, 0.8V and 0.5V). It has standardized 1.2V outputs, and can be seamlessly interfaced with other proposed/conventional pipelines. These significant attributes are collectively achieved by incorporating a DVS Power Unit (embodying power switches), a Local Controller (embodying an asynchronous C-Muller circuit) to execute the pipeline operation, a MUTEX (an asynchronous building block) to synchronize the output, a dual-rail critical path to monitor the completion timing, and buffers/latches with level shifters.

To validate the proposed pipeline, we design a 128-point 16-bit radix-2 Fast Fourier Transform (FFT) processor whose

multiplier is realized based on the proposed pipeline. On the basis of simulations (@ 65nm CMOS), we show that the multiplier based on the proposed pipeline, on average, is $1.94\times$ more power-efficient than that based on the conventional pipeline (within 1MHz to 150MHz).

The paper is organized as follows. Sections II and III describe our proposed pipeline and its critical building blocks respectively. Section IV shows the FFT test vehicle and provides results, and finally conclusions are given in Section V.

II. PROPOSED FINE-GRAINED DVS PIPELINE

Fig. 1 depicts a typical synchronous design, showing three conventional pipelines. Each conventional pipeline has a combinational logic to compute its functional operation, and flip-flops (FFs) to synchronize its output timed by a clock CLK . To enable fine-grained DVS within a pipeline to trade-off speed $\leftrightarrow$ power, we propose a fine-grained DVS pipeline whose block diagram is depicted in Fig. 2. The proposed pipeline is powered a single- V_{DD} (@ V_{DD}) and embodies two power domains, the nominal power domain (shown in the solid bold lines) and the DVS power domain (shown in the dotted bold lines). The nominal power domain is powered by V_{DD} . Inside the nominal power domain, there are buffers, latches, and a DVS Power Unit to provide a secondary power $V-V_{DD}$ to the DVS power domain. The buffers and latches, when necessary, are pre-driven by level shifters to convert signal voltage from $V-V_{DD}$ to V_{DD} , hence standardizing the voltage interface among neighboring pipelines. Put simply, the proposed pipeline can be easily interfaced with other proposed or conventional pipelines. The DVS power domain contains a Local Controller, a combinational logic (consisting of single-/dual-rail logic gates), a MUTEX, and FFs.

The proposed pipeline executes two alternate tasks, i.e. evaluation and reset, for a complete operation. The evaluation

Fig. 1 A synchronous design, showing three pipelines

Fig. 2 Block diagram of the proposed fine-grained DVS pipeline

and reset tasks are in part inferred by the rising-edge and falling-edge CLK respectively. To initiate the evaluation/reset tasks, the handshake signal CLK_IN is triggered, which can be driven by either CLK_OUT of the preceding proposed pipeline or CLK (if a conventional pipeline is used).

For executing an evaluation task, CLK_IN and EN (enable signal) are asserted to '1' so that the Local Controller can generate REQ to trigger the critical path (of the combinational logic). The completion detection signal CD is then generated after the combinational logic (in all critical/non-critical paths) has finished its computation. We note that CD_LS (buffered by a level shifter) is generated to assert the DVS Power Unit to automatically scale $V-V_{DD}$ according to the operating workload (see later). The MUTEX helps resolve the timing issues so that the immediate outputs are stored in the FFs before transferring to the Latches. For executing a reset task, CD_MUTEX (from the MUTEX) feedbacks to the Local Controller to reset the critical path.

Figs. 3 (a)-(d) depict four possible scenarios of DVS for our proposed pipeline; the design of the DVS Power Unit to accomplish these scenarios will be explained in Section III. Fig. 3 (a) is the low power scenario when CLK is slow. Particularly, the assertion of CD is within the logic '1' of CLK , and the negation of CD is within the logic '0' of CLK . In this scenario, $V-V_{DD}$ will drift to the minimum voltage (Min V_{DD}) pre-determined in the DVS Power Unit. Fig. 3 (b) depicts the scenario when the evaluation is slow. Particularly, the assertion and negation of CD is within the logic '0' of CLK . Referring to the 1st operation in Fig. 3 (b), when CLK is switched from '1' to '0' where CD is yet asserted (i.e. slow evaluation), the voltage of $V-V_{DD}$ is scaled-up to speed up the evaluation task, aiming to complete it faster (e.g. within one clock). Fig. 3 (c) depicts the scenario when the reset is slow. Particularly, the negation of CD spans over the next clock. Referring to the Fig. 3 (c), when CLK is switched from '0' to '1' in the next clock where CD is yet negated (i.e. slow reset), the voltage of $V-V_{DD}$ is scaled-up to speed up the reset task, and the degree of voltage charging is stronger than that of in Fig. 3 (b). In this scenario, we define "speculation window" as a time interval between the rising edges of CLK and REQ (in the next clock). As long as the speculation window is within the half-clock duration (assuming 50:50 duty cycle), the proposed pipeline is still error-free. This is because the MUTEX is employed to separate the data capturing (in the FFs) and data outputting (in the Latches), eliminating the interference from the current pipeline to the successive pipeline. This is why we attribute our proposed pipeline to be 'half-clock-tolerant'. Fig. 3 (d) depicts the scenario when both the evaluation and reset are slow. The DVS operation principle is the same as that in Figs. 3 (b) and (c).

For completeness, besides the timing constraint of the speculation window, there are two more timing constraints in our proposed pipeline. The first one is that CD must be asserted within one clock period – this condition is conceptually the same in the conventional pipeline where the critical path delay must be faster than a clock period. The second condition is that when the proposed pipeline is

Fig. 3 Scenarios for DVS (a) slow clock rate (fast evaluation and fast reset), (b) slow evaluation, (c) slow reset, and (d) slow evaluation and slow reset

intermixed with a connecting conventional pipeline, the timing of the speculation window must be accounted into the conventional pipeline. This condition can be easily accounted for by imposing a half-clock-period to the connecting conventional pipeline during synthesis.

III. CRITICAL BUILDING BLOCKS

In this section, the design details of the critical building blocks of the proposed pipeline are discussed.

A. DVS Power Unit

Fig. 4 depicts the DVS Power Unit that is integrated into each pipeline and is designed to adaptively scale $V-V_{DD}$ according to the operating speed of CLK and CD_{LS} . Based on the chosen 65nm CMOS process, the DVS Power Unit is designed to feature three voltage scales via three power switch paths, i.e. the clamped, first charging, and second charging paths. The clamped path (with three NMOS transistors in series) sets the $Min V_{DD}$. In our case, we set the $Min V_{DD}$ to $\sim 0.5V$. The first charging path consists of an NMOS transistor and a PMOS transistor, and is able to drive $V-V_{DD}$ to $\sim 0.8V$. The first charging path will be turned on when the evaluation is slow (see scenarios Fig. 3 (b) and (d)). The second charging path consists of a PMOS transistor, and is able to drive $V-V_{DD}$ to $\sim 1.2V$. The second charging path will be turned on when the reset is slow (see scenarios Fig. 3 (c) and (d)). The sizing of the NMOS and PMOS transistors in these paths depends on the operating workload of the pipeline.

B. Critical Path (Dual-rail gates)

We first identify the critical path from the synthesized netlist based on the Synopsis Design Compiler. Thereafter, we convert the gates in the critical path from the usual single-rail gates into our custom-designed dual-rail differential cascaded voltage swing logic (DCVSL) gates [11]. The single-rail to dual-rail signal conversion [11] for some signals (from non-

Fig. 4 DVS Power Unit

Fig. 5 Local Controller

critical paths) are required in order to drive DCVSL gates. The adoption of the DCVSL gates is to enable a self-timing checking in the sense that the DCVSL gates are propagated one by one in the critical path. In this way, the propagation delays of non-critical paths are inferred to be shorter than that of the critical path. As a result, CD can be used to indicate the completion of the evaluation/reset tasks for the combinational logic. The self-timed DCVSL gates also help to accommodate the Process-Voltage-Temperature (PVT) variations. In short, the tracking of the critical path delay is based on the worst-case, and accounts for the PVT variations.

C. Local Controller and MUX

Fig. 5 depicts the Local Controller, generating REQ on the basis of $NRST$, CLK_{IN} , EN and CD_{MUX} . REQ can only be set to '1' through the primary input (CLK_{IN}) change of the C-Muller circuit (dotted box). However, REQ can only be reset to '0' through the reset function (negative-level sensitive) of the C-Muller circuit. The C-Muller circuit functions as a threshold gate with hysteresis [11]. In overall, the Local Controller serves as a priority arbiter. The top priority goes to the completion of the pipeline. Thereafter, CLK_{IN} can commence a new evaluation by asserting REQ . In short, the set/reset of REQ is alternated in a well-controlled sequence.

Fig. 6 depicts the MUX circuit to exclusively propagate either from CLK to CLK_{MUX} or from CD to CD_{MUX} in the first come first serve basic. Put simply, CLK_{MUX} and CD_{MUX} are never asserted to '1' at the same time. The MUX circuit is used to resolve the timing conflicts. Referring back to Fig. 2, under the operation sequence, CD_{MUX} is used to capture the data in the FFs whereas

CLK_MUTEX is used to latch the data in the Latches and to generate CLK_OUT to indicate the availability of the data.

IV. TEST VEHICLE: A PROPOSED MULTIPLIER PIPELINE IN A FAST FOURIER TRANSFORM (FFT) PROCESSOR

We validate the proposed pipeline by implementing a 128-point 16-bit radix-2 FFT processor. Fig. 7 (a) depicts the pipeline blocks of the FFT processor. Particularly, the multiplier is designed based on our proposed pipeline for DVS because the multiplier often dissipates high power; the power overheads in the proposed pipeline will be easily offset by DVS. The other pipelines remain as the conventional pipelines. The FFT processor is implemented in a 65nm CMOS, and simulated with spice models at $V_{DD} = 1.2V$. Fig. 7 (b) depicts the layout view of the FFT processor.

The FFT processor test vehicle is specifically designed to accommodate a wide operation frequency range from 1MHz to 150MHz, suitable for most of the low-to-mid speed applications. We specially characterize the power dissipation of the multiplier pipeline (with and without the DVS) and their normalized power reading is tabulated in Table I. For clarity, the readings are normalized to that of the proposed pipeline, and the simulated power dissipation values are shown in the parentheses. The conventional pipeline dissipates $1.54\times - 2.47\times$ more power than that of the proposed pipeline. On average, the propose pipeline is $1.94\times$ more power-efficient than the conventional pipeline. It is also interesting to note that at high frequency range (e.g. $\geq 100MHz$), the proposed pipeline has a reduced power-efficiency (compared to that at low clock frequency). It is largely due to frequent charging/discharging of $V-V_{DD}$ to meet the speed requirement when the timing constraints become more stringent.

For more discussion, although both the multipliers embodying the proposed and conventional pipelines are synthesized at 150MHz, the latter can be synthesized at higher

Fig. 6 MUX

Fig. 7 FFT processor: (a) Block diagram and (b) layout view

speed (as it does not embody much overhead). Nonetheless, the proposed pipeline remains advantageous in many aspects because it is power-efficient for low power low-to-mid speed applications, features simple implementation (due to single- V_{DD}), and is adaptable within the existing synchronous design methodology. For higher power efficiency, the multi- V_{DD} could be explored in the future.

Table I
Comparison: Power dissipation ($\mu W @ V_{DD} = 1.2V$) of the multiplier embodying the proposed DVS pipeline and conventional pipeline

Technique Frequency	Conventional Pipeline (without DVS)	Proposed fine- grained DVS Pipeline
1MHz	$2.26\times$	$1\times (24.3)$
10MHz	$2.47\times$	$1\times (40.0)$
25MHz	$2.07\times$	$1\times (88.0)$
50MHz	$1.85\times$	$1\times (169.0)$
75MHz	$1.80\times$	$1\times (246.0)$
100MHz	$1.61\times$	$1\times (355.0)$
150MHz	$1.54\times$	$1\times (538.0)$
Average	$1.94\times$	$1\times$

V. CONCLUSIONS

We have proposed the DVS pipeline, featuring fine-grained DVS which trades off speed for power innately (i.e. default scenario when speed is inconsequential) or power for speed (when speed is beneath the requirement). We have also shown that the proposed DVS pipeline supports fast-transition and remains error-free up to one-and-a-half clock duration per operation. The power-efficiency of the proposed DVS pipeline has been validated through a FFT processor ($1.94\times$ lower power compared to conventional pipeline). We recommended it to be applied for low power low-to-mid speed applications.

REFERENCES

- [1] T. Lin, *et al.*, "An Ultra-Low Power Asynchronous-Logic In-Situ Self-Adaptive V_{DD} System for Wireless Sensor Networks," *IEEE JSSC*, v48, pp. 573-586, 2013.
- [2] Z. Kejiu, *et al.*, "New Insights on Dynamic Voltage Scaling of Multiphase Synchronous Buck Converter: A Comprehensive Design Consideration," *IEEE Trans. Power Electronics*, v29, pp. 1927-1940, 2014.
- [3] H. W. Lee, *et al.*, "A 1.6 V 1.4 Gbp/s/pin Consumer DRAM With Self-Dynamic Voltage Scaling Technique in 44 nm CMOS Technology," *IEEE JSSC*, v47, pp. 131-140, 2012.
- [4] Y. Shakhsher, *et al.*, "A 90nm data flow processor demonstrating fine grained DVS for energy efficient operation from 0.25V to 1.2V," *IEEE CICC*, 2011, pp. 1-4.
- [5] M. Putic, *et al.*, "Panoptic DVS: A fine-grained dynamic voltage scaling framework for energy scalable CMOS design," *IEEE ICCD*, 2009, pp. 491-497.
- [6] A. Rahimi, *et al.*, "Dynamic voltage scaling for fully asynchronous NoCs using FIFO threshold levels," *IEEE CADs*, 2010, pp. 43-48.
- [7] S. Das, *et al.*, "A self-tuning DVS processor using delay-error detection and correction," *IEEE JSSC*, v41, pp. 792-804, 2006.
- [8] M. Nakai, *et al.*, "Dynamic voltage and frequency management for a low-power embedded microprocessor," *IEEE JSSC*, v40, pp. 28-35, 2005.
- [9] T. D. Burd, *et al.*, "A dynamic voltage scaled microprocessor system," *IEEE JSSC*, v35, pp. 1571-1580, 2000.
- [10] T. Kuroda, *et al.*, "Variable supply-voltage scheme for low-power high-speed CMOS digital design," *IEEE JSSC*, v33, pp. 454-462, 1998.
- [11] R. Zhou, *et al.*, "A Low Overhead Quasi-Delay-Insensitive (QDI) Asynchronous Data Path Synthesis Based on Microcell-Interleaving Genetic Algorithm (MIGA)," *IEEE TCAD*, vol. 33, no. 7, pp. 989-1002, 2014.