Quasi-Resonant Clocking: Continuous Voltage-Frequency Scalable Resonant Clocking System for Dynamic Voltage-Frequency Scaling Systems

Fahim ur Rahman, Student Member, IEEE, and Visvesh Sathe, Member, IEEE

Abstract—This paper describes quasi-resonant-clocking (QRC), a continuous voltage-frequency scalable resonant clocking architecture capable of dynamic voltage and frequency scaling (DVFS). The use of runtime control by QRC is central to ensuring robust, efficient operation across variations in process, voltage, and temperature. QRC exhibits instantaneous wide-range duty-cycle control, required for a broader class of clocking applications involving large capacitive loads. Two QRC variants, QRCfoot and QRCpass, are presented with test-chip measurements under DVFS operation in 65-nm CMOS. The QRC test-chip demonstrates a maximum energy reduction of 47% at 132 MHz. Across both implementations, energy measurements based on runtime DVFS in 0.7–1.2 V range indicate energy reduction in the 32%–47% range.

Index Terms—Dynamic voltage and frequency scaling (DVFS), low power, process-voltage-temperature (PVT) tolerant, resonant clocking, runtime delay-locked-loop (DLL)-based control, system clocking, tunable duty-cycle, voltage-frequency-scalability.

I. INTRODUCTION

POWER dissipation continues to play a central role in impacting computing performance, mobility, and cost over a broad range of digital systems, from high-performance microprocessors [1]–[3] to ultra-low-power systems employing aggressive pipelining [4]. Fine-grained clock gating has significantly reduced clock power [3], [5], but the large global clock distribution driving these clock-gates accounts for a sizable fraction of total system power [6], [7]. Enabling clock power reduction, therefore, remains critical to the efficient implementation of digital systems.

Resonant clocking has been proposed as an efficient approach to global clock distribution [8]–[19]. Fig. 1 illustrates the central idea behind resonant-clocking—introducing inductance (L) into the clock network, and enabling efficient LC resonance between the clock load (Cload) at frequencies close to the natural (or resonant) frequency, \( f_0 = 1/(2\pi LC^{1/2}) \).

In contrast with conventional \( CV^2 \) per-cycle energy dissipation, the energy delivered by the supply to a resonant clock system each cycle needs to only compensate for the \( I^2 R \) losses in the LC tank. High quality-factor (Q) designs, achieved through reduced \( R_L \) and \( R_{clk} \) sustain clock oscillations with a significantly lower energy dissipation per cycle (\( E_{PC} \)): \( E_{PC} = \pi/4QC_{load}V_{dd}^2 \) [15].

Recent implementations have overcome several outstanding processor integration challenges. Two such important challenges are the design of inductors in-line with power trunks [6], [7], [14], [20]–[22], and addressing system-Q degradation due to power grid eddy-current flow [6]. The result of these efforts has been processor designs with significant clock power reduction, positively impacting power-constrained performance and battery-life.

Meanwhile, the increased significance of dynamic voltage and frequency scaling (DVFS) [23], and a trend toward its...
increasingly aggressive use across a broader voltage-frequency range dilutes resonant clocking efficiency benefits. Resonant clock efficiencies, attractive at clock frequencies close to $f_0$ (Fig. 1), quickly diminish along with $Z_{\text{tank}}$ outside a small frequency range away from $f_0$. Although power-constrained performance benefits remain, the broader energy-efficiency advantages of the technique are diminished as systems spend a smaller portion of time within this narrow frequency range.

Recent works have attempted to address this challenge using techniques suitable for their specific applications [7], [24]. However, these techniques face challenges of limited frequency-tunability [7], limited duty cycle control [24], or clock waveforms that transitions to $-V_{\text{dd}}$, limiting scalability [24].

In this paper, we elaborate upon quasi-resonant-clocking (QRC) [25], a DVFS-compliant technique that achieves continuous voltage and frequency scalability. The proposed approach produces rail-to-rail clock waveforms with near-arbitrary duty-cycle control. With the exception of [24] (itself restricted to sub-threshold designs) the proposed approach is the only architecture capable of achieving uniform energy efficiency continuously across the 0 to $f_0$ frequency range. A key enabler of efficient, process-voltage-temperature (PVT) robust operation, and DVFS support is autonomous timing control of QRC related circuits which can track slewing supply-voltages during DVFS events, and temperature variations. The proposed technology is capable of a frequency range of 0–$f_0$, and sub-threshold to nominal-$V_{\text{dd}}$ operation. We present circuits and architectures for two different variants of QRC, notably QRC$^{\text{pass}}$ and QRC$^{\text{foot}}$, which offer different capabilities and efficiencies. We also present an analysis of QRC energy and its dependence on key QRC parameters. This model is used to determine optimal design parameters, and examine the fundamental tradeoff between efficiency and rise/fall times exhibited by all resonant-clock systems. The two proposed QRC variants are validated through test-chip demonstrations of a DVFS system consisting of an eight-way multiply-accumulate (MAC) array in 65-nm CMOS. Last, we present test-chip measurements obtained from each of these two implementations.

The remainder of this paper is organized as follows. In Section II, we review related work in the area of frequency-scalable resonant clocking. In Section III, we provide an overview of the QRC architecture and the salient aspects of its operation. Circuit and architecture descriptions for each variant of the QRC implementation are discussed in Sections IV and V. An analysis of QRC, including the examination of energy-optimal design parameters is presented in Section VI. Test-chip architecture, evaluations and measurements of both variants, QRC$^{\text{foot}}$ and QRC$^{\text{pass}}$, are presented in Section VII.

II. RELATED WORK

Several prior works in the literature have focused on addressing the limited frequency range of resonant clocking. These approaches range from disabling resonant clocking at frequencies outside a narrow range [6], to alternative resonant clocking implementations which extend resonant operation outside its narrow operating range.

A technique originally proposed and demonstrated in [26], [27] for I/O pad and display drivers, and re-implemented more recently in [28] involves using switched capacitor banks to transition a clock load through a sequence of intermediate voltage levels. Although slightly less efficient than inductor-based techniques, the switched-capacitor-based technique offers notable advantages: 1) simpler logic control; 2) ease of integration; and 3) natural voltage scalability. However, though effective for its originally intended application of driving adiabatic logic or LCD drivers [26], [27], this approach is not well suited to system clocking—employing the technique efficiently results in clock waveform “shelving” [27] (the clock voltage levels off at intermediate voltages during) that impacts efficiency and race-immunity among timing elements. In more severe situations, non-monotonic clock waveforms near mid-rail voltages [28] can result in runt clock pulses.

More recently, intermittent resonant clocking, a novel technique achieving uniform efficiency resonant clocking from 0–$f_0$ has been proposed [24]. Illustrated in Fig. 2(a), the technique presents a novel topology to deliver an $RLC$ oscillation enabled clock “blip” from $V_{\text{dd}}$ to $-V_{\text{dd}}$. The pMOS can hold clk at $V_{\text{dd}}$ for an arbitrary duration before commencing the next “blip.” Frequency scaling is achieved by varying the duration of time that the clk net remains at $V_{\text{dd}}$. The proposed architecture, developed for sub-threshold operation, does not scale to higher voltages by construction. Non-compliant CMOS voltages—the clock waveform transitions below 0 V to a body-drain diode-drop voltage of approximately 700 mV—prevent operation at nominal voltages. Furthermore, operation at
higher voltages also results in the dissipative turn-on of the nMOS diode turn-on as clock transitions below $-700 \text{ mV}$, shunting the RLC system. Finally, the inability to control duty cycle limits the approach to use in flip-flop-based system clocking applications.

Bandwidth extension through additional programmable inductors to modulate total inductance [Fig. 2(b)] and shift $f_0$ has been demonstrated [7], and recently implemented in a production processor [22]. The proposed approach caters to its target application, but does not provide continuous voltage-frequency scalability across a broad range. It is, therefore, not well suited to applications seeking a truly broad resonant-clock operating frequency range. Furthermore, the placement of additional inductors in an already physically and electrically constrained environment places additional inductor design challenges and complexity, and relies on technology to deliver inductors of sufficient quality-factor.

III. QRC OVERVIEW

QRC achieves continuous frequency scalability by effectively interleaving conventional and resonant modes of clock operation. Fig. 3 illustrates the rationale behind the QRC architecture. The time instant when the clock voltage is at its maximum or minimum is ideal for disconnecting the inductor from the clock domain (discussed further in this section). Disconnecting the inductor safely and efficiently enables the clock to be held to the supply and ground rail for an arbitrary duration of time ($\tau_{\text{high}}$ and $\tau_{\text{low}}$, respectively), readily achieving frequency and duty-cycle control.

A simplified circuit diagram for the proposed QRC clocking scheme is shown in Fig. 4(a). Conduction switch $M_c$ conditionally disconnects the inductor from the clock network at the end of each resonant transition, twice every cycle. $C_{\text{acg}}$ blocks dc inductor current flow and is large enough to act as an ac ground connection.

Fig. 4(b) shows a simplified QRC timing diagram, identifying signal transitions for $n$, $p$, and $c$, controlled precisely through the Timing Control module. Consider the steady-state operation of the QRC system. The voltage across $C_{\text{acg}}$ ($C_{\text{acg}} \approx 30 \cdot C_{\text{load}}$) remains steady at $V_{\text{dd}}/2$ (for a 50% duty-cycle clock). At the start of the clock cycle, $V(\text{clk}_{\text{PLL}}) = 0$, $V(\text{clk}) = 0$. As $\text{clk}_{\text{PLL}}$ transitions to $V_{\text{dd}}$, $c$ is asserted and connects the inductor to the resonant system, effecting an RL current build-up in the inductor. $T_{\text{BLD,n}}$ and $T_{\text{BLD,p}}$ correspond to the inductor-current build-up times before the rising and falling transitions of the resonant clock, respectively. Use of $T_{\text{BLD,p}}$ and $T_{\text{BLD,n}}$ are motivated by two factors.

1) A longer build-up provides a sharper slew for the output clocks, essential for reliable digital systems.

2) Suitably tuned build-up durations provide improved efficiency [6], [15]. The optimal build-up duration varies depending on a number of factors including the resistance in the build-up and resonant paths and the load capacitance.

After duration $T_{\text{BLD,n}}$, $n$ transitions to 0, turning off the hitherto conducting $M_n$, starting the LC transition of clk from 0 to $V_{\text{dd}}$. After a delay ($T_{\text{RE}}$) determined by the natural frequency of the LC network, clk reaches its peak voltage. A comparator detects this peak, prompting the timing module to de-assert $c$ and disconnect the inductor from the network precisely when inductor current $I_L = 0$ with help from a timing-control module. Concurrently, $p$ transitions to 0, turning on $M_p$, and driving clk to $V_{\text{dd}}$ rail.
this sequence, clk has transitioned to $V_{dd}$ and the inductor is disconnected from the grid, enabling $M_p$ to hold clk to $V_{dd}$ indefinitely until the next transition of clk toward 0. The sequence of events during the falling clk edge is conceptually identical to the rising edge sequence with a reversal in $I_L$ direction. The occurrence of clk maxima or minima coincides with $I_L = 0$, which also results in a reversal of polarity of the conducting $M_c$ switch. clk maxima or minima detection is therefore performed by comparing the potential difference across $M_c$.

The clock cycle time, $T_{clk}$ can be written as: $T_{clk} = T_{BLD,p} + T_{RE} + T_{BLD,n} + T_{FE} + T_{low}$ Frequency and duty cycle programmability is achieved by varying $T_{high}$ and $T_{low}$ independently in QRC [29]. $T_{RE}$ and $T_{FE}$ are determined by the natural frequency of the $LC$ system: $T_{RE} + T_{FE} = 1 / f_0 = 2 \pi L C^{1/2}$. Consequently, the maximum operating frequency using QRC is therefore limited by $f_0$, corresponding to ordinary resonant operation.

DVFS events require charging $C_{acg}$ to $V_{dd}/2$. However, at prevalent levels of capacitance ($\approx 30 \cdot C_{load}$), the charge-discharge durations amount to a few clock cycles, and do not pose an overhead for DVFS. In addition, any settling time resulting from the voltage transition has a minor temporary impact on clock slew and efficiency.

To provide PVT-robust, efficient clocking compatible with runtime voltage and frequency transitions, QRC faces two important challenges–timing control, and the efficiency of the $M_c$ switch. Turning $M_c$ off while $I_L \neq 0$ results in wasteful dissipation and voltage ringing that exceeds the supply rail, degrading both efficiency and reliability. Furthermore, precise timing must be maintained across variation in PVT, and across the DVFS-enabled $V_{dd}$ range. An all-digital runtime control system relying on a sense-amplifier and dual-edge triggered delay-locked-loop (DLL) provides the necessary current detection.

The second challenge of ensuring robust and efficient operation arises from the quiescent $V_{dd}/2$ voltage across $C_{acg}$, motivating design of a gate driver to meet the resulting gate-overdrive needs of $M_c$. These challenges manifest separately in QRC_foo and QRC_pas and are therefore addressed individually through different gate-driver circuits.

IV. QRC_foo Design

In this section, we discuss the architecture and circuits that constitute the QRC_foo 65 nm test-chip implementation. In particular, we describe two key enabling sub-systems. A DLL-enabled Timing Control module and the footer-driver circuit.

Fig. 5 illustrates the QRC_foo architecture [25]. $M_c$ is connected as a footer device between the bottom-plate of $C_{acg}$ and ground. A footer arrangement provides full gate-overdrive ($V_{gs} = V_{dd}$). The Timing-control module uses a comparator sampling the source-drain terminals of $M_c$, and clocked by $c$ to provide feedback to the Timing Control module to ensure carefully timed assertions and de-assertions of $p$, $n$, and $c_{pre}$. Although $M_c$ enjoys full gate-overdrive when $V = V_{dd}$, a driver circuit is required enforcing device cutoff when $V_{clk}$ is in conventional mode, as discussed further in Section IV-B.

A. Timing Control Module

The Timing Control module orchestrates the assertion/de-assertion of $n$, $p$, and $c$ to enable robust, efficient operation under slewing $V_{dd}$ and $f_{clk}$ conditions expected in DVFS. Timing requirements for $n$, $p$, and $c$ are stringent to ensure the inductor is disconnected from the clock domain when $I_L = 0$ at the end of each resonant transition to avoid wasteful, reliability-degrading under-damped voltage oscillations. The assertion and de-assertion delays of $n$ and $p$ are determined by frequency of the output clock ($f_{clk}$, $f_0$, $T_{BLD,n}$, $T_{BLD,p}$) (Fig. 5). Finally $c$ switches twice every cycle, and is asserted for durations $T_{RE}$ and $T_{FE}$ of each resonant transition of clk.

Fig. 6 summarizes the implementation of the Timing Module. A DLL implementation aligns two events—the de-assertion of $c$ (the gate-driver output, which disconnects the inductor from the network), and the advent of zero current flow in the inductor. The need to align these events twice every cycle (one for each clock edge) when transitioning from resonant to conventional mode motivates a “dual-DLL” architecture, with a shared front-end phase-detector and two delay-chains to achieve the desired timing control for each transition. Independent delay chains provide separate control of current buildup durations ($T_{BLD,n}$, $T_{BLD,p}$) and clock transition times ($T_{RE}$, $T_{FE}$) (Fig. 5) for rise and fall clock edges to support non 50% duty-cycles.

A p-type strong-ARM latch, triggered on the de-asserting (for an nMOS $M_c$) edge of $c$, performs
phase-detection, determining current flow direction through \( M_c \) by sampling the relative polarity of its source and drain terminals twice every cycle. Thus, the phase-detector determines whether \( c \) is de-asserted too early or late relative to the \( I_L = 0 \) event based on the direction of current flow (with \( C_{acg} \) either sourcing or sinking current). This phase-detector output provides the required early/late signals to the dual-DLL. The DLL controller updates the rise and fall delay chains, with the \( \text{rise} \_\text{delay} \) (\( \text{fall} \_\text{delay} \)) chain updated based on the comparator sample when the clk voltage is \( V_{dd} \) (0). Delay-code update-driven glitches are avoided by updating delays when delay-chain outputs are not observed—Updates to the \( \text{rise} \_\text{delay} \) (\( \text{fall} \_\text{delay} \)) chain occur when clk is 0 (\( V_{dd} \)). Glitch-less, context-dependent delay for \( n \), \( p \), and \( c \) is provided by using selection MUXes.

By starting resonant clock transitions a fixed delay (\( T_{BLD,n} \) and \( T_{BLD,p} \)) after clkPLL, QRC is transparent to duty-cycle changes made at its input, on clkPLL. This feature is beneficial for optimizing phase-paths in some digital designs. Furthermore, any changes in input clock duty-cycle are instantaneously referred to the output, a key benefit in IVR applications involving runtime modulation of clock duty-cycle.

Programmable delay is achieved through a telescopic delay chain [30] for coarse resolution, combined with a programmable load inverter chain for fine delay control (Fig. 7). A thermometer code controls the coarse telescopic delay chain [30], and MOS-based capacitors offer programmable load in the fine delay chain. The delay chain was designed to provide sufficient resolution at \( V_{dd} = 0.7 \) V, and range at \( V_{dd} = 1.2 \) V for robust, efficient QRC operation.

Fig. 8 summarizes the operation of the DLL controller. Upon cold-start, the system first safely determines PVT-dependent delay-code settings. Reliability degradation due to initially mistimed \( c \) signals upon start-up are managed by first locking with only a fraction of the \( M_c \) device, limiting \( I_L \), and throttling any under-damped voltage oscillations. After achieving DLL lock for both edges (rising and falling), the controller increases \( M_c \) width and returns to locking mode, making any necessary delay adjustments to re-lock. This interleaved lock and \( M_c \)-width update is repeated until the target \( M_c \) width is reached. Once locked, the DLL transitions into a low-power mode, with a sampling rate (kilohertz range) required tracking temperature changes on the die, rendering DLL controller power negligible.

**B. Footer Driver Circuit**

The QRCFooter implementation employs a footer for efficient operation through full gate overdrive. However, when \( V(clk) = 0 \), the \( V_{dd}/2 \) voltage across the \( C_{acg} \) results in the footer drain voltage reaching \( -V_{dd}/2 \) (Fig. 9). Using traditional drivers to set the gate voltage to 0 V is therefore insufficient, and results in current back-flow from the ground terminal which discharges \( C_{acg} \). A footer driver is therefore designed to drive \( c \) to achieve full gate overdrive when on, and ensure cut-off when non-conducting. When the network is in resonant mode with \( V(c_{pre}) = V_{dd} \), the driver drives \( c \) to \( V_{dd} \), allowing \( M_c \) to conduct in the linear mode regardless of clock polarity. When clk is held at \( V_{dd} \) in conventional mode with \( V(c_{pre}) = 0 \), the drain voltage of \( M_c \) (\( n_{acg,bot} \) transitions to \( V_{dd}/2 \)) sets the driver drives \( c \) to 0 V. However, when \( V(c_{clk}) = 0 \) in conventional mode, the drain of \( M_c \) transitions to \( -V_{dd}/2 \), requiring that the driver set \( c \) to \(-V_{dd}/2 \) to ensure \( M_c \) is in cut-off.

We implemented a footer driver [Fig. 9(b)] to efficiently enable this functionality. The pull-up network consisting of \( M_0 \) and \( M_1 \) sets \( c \) to \( V_{dd} \) during build-up and resonant transition. In conventional mode, disabling \( M_c \) when \( V(c_{clkPLL}) = V_{dd} \) similarly involves a conventional pull-down through conducting devices \( M_2 \) and \( M_3 \), \( M_4 \) and \( M_5 \) are in cutoff and do not interfere in the action of the pull-down network. When \( V(c_{clkPLL}) = 0 \) in conventional mode, \( n_{acg,bot} \), the erstwhile drain of \( M_c \) is \(-V_{dd}/2 \), \( M_5 \) conducts to ensure that \( c \) is connected to \( n \), ensuring \( V_{gs} = 0 \). \( M_4 \) is employed to connect the gate terminal of \( M_2 \) to \( n \), thereby avoiding current back-flow in the pull-down stack. Thick oxide devices are required for \( M_4 \) and \( M_5 \) to withstand higher oxide stress when

---

**Fig. 7.** Delay chain design ensures the range and precision required at across the entire \( V_{dd} \) range. (a) Coarse delay chain. (b) Fine delay chain.

**Fig. 8.** DLL lock acquisition procedure, avoiding reliability challenges before DLL-lock, and finally transitioning to a low-bandwidth mode sufficient to track temperature variation.
conducting. The higher $V_{th}$ of these devices is offset by the increased gate overdrive of $3/2V_{dd} - V_{th}$ during conduction. $M_1$ and $M_7$ are protection devices employed to prevent oxide stress related degradation in $M_0$ and $M_6$, respectively.

V. QRCpass Design

In this section, we discuss circuits and architectures that constitute the QRCpass design. The QRCpass topology enables a single inductor to be shared between multiple clock load domains for a Single Inductor multiple output capability. This feature is promising for a variety of applications including driving multi-phase bridges of an integrated buck converter. Similar to QRCfoot, QRCpass relies on the same Timing Control module for precise signal timing. In this section, we focus on modules that are distinct from the QRCfoot implementation, namely, the sense-amplifier and the gate-overdrive circuits.

QRCpass employs a gate-boosted nMOS pass-gate ($M_c$) to serve as the conduction switch (Fig. 10), resulting in a set of current-sensing and gate-overdrive challenges distinct from those in QRCfoot. Unlike QRCfoot, where the sense-amplifier measures $M_c$ source–drain polarity around a 0 V common mode, QRCpass requires polarity sensing with a common-mode of $V_{dd}$ and 0 V after the rising and falling edges of clk, respectively. In addition, the $V_{dd}/2$ voltage across $C_{acg}$ results in reduced gate overdrive for $M_c$.

A. Sense-Amplifier Latch

Fig. 11 shows the proposed sense amplifier (sense-amp) implemented for QRCpass. In order to sample twice every cycle, with common mode voltages of 0 and $V_{dd}$, respectively, the proposed sense-amplifier exhibits alternating operation of pull-down and pull-up sense-amplifier sections. When $clk_{PLL} = 1$, the sense-amplifier amplifies the differential current drive in $M_1$ and $M_2$ (enabled by footer device $M_0$) to determine current flow direction in $M_c$. Devices $M_3 - M_6$ provide the positive feedback to resolve the differential current drive, $M_{10} - M_{13}$ serve to pre-charge sense-amplifier nodes during both clock transitions. Finally $M_{15}, M_{17}$ ($M_{14}, M_{16}$) provide the necessary power (ground) connection to the sense-amplifier when $clk = 1$ ($clk = 0$). Devices $M_1 - M_8$ are up-sized, and employ non-minimum channel length devices to mitigate random mismatch between differential devices.

B. Gate Boosting

Placing $M_c$ between the inductor and clk allows multiple domains to share a single inductor, but results in degraded gate overdrive. A standard 2X voltage doubler [31]–[33] applied to $M_c$ provides the necessary gate overdrive but is not feasible due to the $2V_{dd}$ gate–source voltage that will result as clk approaches 0. We propose a gate boosting topology suited to the QRCpass implementation [Fig. 12(a)] that relies on bootstrapping the gate of $M_c$ with the transition of $clk_{ind-side}$. The resulting system provides full gate overdrive ($V_{dd} - V_{th}$) for $M_c$ throughout the resonant clk transition between the supply rails. The proposed circuit ensures that e the gate signal provided to $M_c$ is 0 V when $c_{pre} = 0$. When $c_{pre} = 1$, the boosting circuit delivers a voltage of $1.2V + V_S$ (where $V_S$ is the source voltage of $M_c$), consistently ensuring full gate-overdrive for $M_c$.
and gate–drain voltage during operation, the gate voltage and falling. Notably, while maintaining a 1.2-V gate–source maintained during the entire resonant transition, both rising and conductivity of the pass-gate transitions from either $V_{dd}$, resulting in a voltage doubler gate-oxide stress compliant of the pass-gate boosting. The proposed topology provides a constant 1.2-V $V_{gs}$ to $M_c$ while using only logic devices while maintaining reliability. (b) Pass-gate boosting simulation waveforms.

Fig. 12(b) illustrates the operation of the proposed circuit. In conventional mode ($c_{pre} = 0$), series devices $M_1$ and $M_2$ (a gate-oxide protection device) drive $c$ to 0 V, ensuring cutoff. The boosting capacitor ($C1$) pre-charges to $1.2$ V. In resonant mode, as $c_{pre} = 1$, turning $M_c$, $TX$ conducts ($c_{pre} = 1$) connecting the bottom-plate of C to $n_{Lside}$. $M_3$ conducts, driving out to $V(n_{Lside}) + 1.2$ V. This gate overdrive is maintained during the entire resonant transition, both rising and falling. Notably, while maintaining a 1.2-V gate-source and gate-drain voltage during operation, the gate voltage safely transitions up to $2.4$ V (When $V_{dd} = 1.2$ V) while $M_c$ conducts. The gate-oxide of $M_c$ is protected by the conducting channel while conducting.

An off-chip 1.2-V supply voltage was used to aid test and characterization of the proposed circuit. An integrated implementation of the supply using charge-pumps is relatively straightforward due to the low-current draw of the 1.2-V supply. Alternatively, the 1.2-V supply can be replaced with $V_{dd}$ resulting in a voltage doubler gate-oxide stress compliant design at the cost of reduced efficiency at low voltages.

In QRC_pass, turning off of the pass gate causes an under-damped RLC oscillation at node $n_{sw}$ as the drain capacitance of the pass-gate transitions from either $V_{dd}$ or 0 to $V_{dd}/2$ [Fig. 13(b)]. This underdamped oscillation is unavoidable, but does not affect the reliability of devices in the QRC system.

Fig. 12. (a) Schematic circuit for pass-gate boosting. The proposed topology provides a constant 1.2-V $V_{gs}$ to $M_c$ while using only logic devices while maintaining reliability. (b) Pass-gate boosting simulation waveforms.

VI. ANALYSIS

A. Energy Dissipation Analysis

In this section, we present simulations and analytical models for QRC energy dissipation consisting of switching and conduction loss components. Simulations are used to examine aspects of QRC design that are either onerous or impossible to demonstrate through measurement of the test-chip. Given the similarity of the analysis between QRC_pass and QRC_foot, we restrict our focus on the QRC_pass implementation for brevity. We use the simplified schematic in Fig. 13(a) to analyze the energetics of QRC.

The energy-per-cycle (EPC) dissipation of a resonant clocking system modeled as an equivalent $R$, $L$, and $C$ system can be approximated as

$$E_{PC_{res}} = \frac{\pi}{4Q}C V_{dd}^2$$

where $Q$ is the system quality factor.

To analyze QRC energy dissipation, we start by noting that a lossless LC resonant clock transition is sinusoidal, with magnitude $V_{dd}/2$ around a dc voltage level $V_{dd}/2$. The total capacitance $C_{total} = C_{load} + C_p$ stores a charge of $C_{total}V_{dd}/2$, corresponding to an energy storage of $1/2C_{total}(V_{dd}/2)^2$.

In practice, $C_{load}$ charges to a lower value, $V_{peak}$ [Fig. 13(b)], the difference accounting for the conduction losses in the system. $V_{peak}$, the voltage of clk at the end of its resonant transition, and $E_{cond}$, the per-cycle conduction loss can be determined using known values of $Q$ and

$$V_{peak} = \frac{V_{dd}}{2}(1 + e^{-\frac{\pi}{\sqrt{4Q^2 - 1}}})$$

$$E_{cond} = (C_{load} + C_p) V_{dd}(V_{dd} - V_{peak}) = (C_{load} + C_p) V_{dd}[V_{dd} - \frac{V_{peak}}{2} - (1 + e^{-\frac{\pi}{\sqrt{4Q^2 - 1}}})].$$

An additional source of energy loss arises from conventional charge and discharge of parasitic capacitance, $C_p$ on the inductor side of the switch. Twice every cycle, this capacitance transitions from $V_{dd}/2$ to the supply rails at the onset of the resonant transition and returns to the mid-rail voltage after the transition. The resulting energy dissipation, $E_{sw,parasitic}$ can be modeled as

$$E_{sw,parasitic} = 4 \cdot \frac{1}{2} C_p V_{dd}^2 = \frac{1}{2} C_p V_{dd}^2.$$  

Similarly, the system incurs switching losses in the driver twice, once for each transition

$$E_{sw,driver} = 2 \cdot \frac{1}{\eta_{boost}} \cdot C_M c \cdot (V_{dd} + 1.2)^2$$

where $C_M c$ is the capacitance of the $M_c$ switch and $\eta_{boost}$ is the efficiency of the gate boosting module. Adding (2)–(4) provides the expression for total losses per-cycle

$$E_{tot} = (C_{load} + C_p) V_{dd}(V_{dd} - V_{peak}) + \frac{1}{2} C_p V_{dd}^2$$

$$+ \frac{2}{\eta_{boost}} \cdot C_M c \cdot (V_{dd} + 1.2)^2.$$  

The tradeoff between switching and conduction losses is apparent in $M_c$ width selection (Fig. 14). Increasing $M_c$ width reduces resistance, and $E_{cond}$, at the expense of increased $E_{sw,driver}$. Including the resistive contribution of $M_c$ in the $I^2R$ losses in (2), and using the resulting $V_{peak}$
to obtain \( E_{\text{tot}} \), the resulting analytical EPC results derived from (3)–(5) match well with simulation results (with parasitic capacitance).

Fig. 14 shows the existence of an optimal \( M_c \) width, trading off switching and conduction losses in (5). As seen from (2), increased resistance due to the clock distribution, the inductor, or reduced series switch width, degrades system Q, increasing conduction losses. On the other hand, larger switches increase switching losses. In our proposed implementation of QRC\(_{\text{pass}}\) and QRC\(_{\text{foot}}\), the switch-width was designed to be tunable by having several banks of MOSFET in parallel. In post-silicon, a one-time calibration was run to fix the optimal switch width.

**B. \( f_{\text{max}} \)**

For QRC, clock frequency is modulated by inserting a hold-high/hold-low duration between the resonant transition and changing the duration of them. The maximum frequency is possible when \( \tau_{\text{hold high}} = \tau_{\text{hold low}} = \tau_{\text{buildup}} = 0 \), which corresponds to a sinusoidal clock waveform operating at its natural frequency, \( f_0 \).

Increased \( f_0 \) delivers broader operating range but at the expense of reduced energy savings (1)

\[
\text{EPC}_{\text{res}} = \frac{\pi}{4Q}CV_{dd}^2 = \frac{\pi}{4} (2\pi f_0 RC) CV_{dd}^2 = \text{EPC}_{\text{conv}} \left( \frac{1}{2} \pi^2 f_0 RC \right).
\]

\( f_{\text{max}} \) can be calculated from the targeted energy savings

\[
\text{Energy efficiency, } \eta_{\text{res}} = 1 - \frac{\text{EPC}_{\text{res}}}{\text{EPC}_{\text{conv}}} = f_0 = \frac{2(1 - \eta_{\text{res}})}{\pi^2 RC}.
\]

From (7) it can be seen that increasing \( f_0 \) degrades efficiency. This trend imposes a tradeoff between energy-efficiency and maximum achievable frequency. In designs with extremely high clock loading (global clock distributions of microprocessors), multiple distributed inductors have been shown to be effective, with each inductor resonating with a portion of the total clock load (see [6], [7]).

**VII. TEST-CHIP ARCHITECTURE AND MEASUREMENTS**

In this section, we present the architecture and implementation of the QRC test-chip featuring the QRC\(_{\text{foot}}\) and QRC\(_{\text{pass}}\) variants, and their test-chip measurements.

**A. Test-Chip Architecture**

QRC\(_{\text{foot}}\) and QRC\(_{\text{pass}}\) are implemented separately as DVFS compliant clock distributions for an eight-way MAC array datapath in 65-nm CMOS (Fig. 15). Both datapaths employ build-in self test and scan-chains for functional and parametric test. The eight-way MAC was implemented using synthesis, auto-place and route (SAPR). The SAPR flow was augmented to produce a tapered H-Tree wire clock distribution driving a three-level H-Tree driving a clock mesh on metal layers M4 and M5. The clocked comparators, footer-driver, charge pump, delay circuits, and the clock drivers were custom developed. Resonant clocks typically exhibit relatively lower skew due to low network resistance [6], [7]. The inherent skew degradation in resonant clocks, however, causes significant PVT-induced skew in post-gater clock distributions, leading to potential setup and hold-time degradation. Of the several techniques available to mitigate this degradation [6], [7], [15], we adopt latch-based design [15] approach. \( C_{\text{acg}} \) is implemented as an off-chip electrolytic capacitor for both variants.
This test-chip used an off-chip inductor in the implementation of QRC\textsubscript{foot} and QRC\textsubscript{pass}. While on-chip inductors can be used for QRC, the choice of off-chip inductors was driven by the desire to explore the impact of inductor selection for slewrate and efficiency tradeoffs, and be able to readily capture resulting QRC clock waveforms (Fig. 16). Notably, use of on-chip inductors does not significantly degrade system efficiency because $M_e$ losses largely dominate system $Q$ ($\approx 1.3$ for QRC\textsubscript{pass}, $\approx 1.2–1.8$ for QRC\textsubscript{foot}). Robust system operation was ensured through the design of the resonant-clock latch-based design methodology [15]. Runtime voltage-frequency scaling was performed by using a ring-oscillator powered by the logic supply as a clock source. The energy measurements reported in this paper include the dissipation from all constituent QRC sub-systems. In QRC\textsubscript{foot} and QRC\textsubscript{pass}, the value of $C_{\text{acg}}$ and $L$ are $5$ nF and $7$ nH, respectively. The extracted value of capacitance, $C_{\text{load}}$ is approximately $180$ pF.

As mentioned in Section III, an energy-optimal selection of $T_{\text{BLD,p}}/T_{\text{BLD,n}}$ exists. In the QRC\textsubscript{pass} and QRC\textsubscript{foot} implementations, $T_{\text{BLD,p}}/T_{\text{BLD,n}}$ was implemented with tunable delay chains. After a one-time calibration, the optimum $T_{\text{BLD,p}}/T_{\text{BLD,n}}$ duration was found to be $80$ ps at $1$ V.

### B. QRC\textsubscript{foot} and QRC\textsubscript{pass} Test-Chip Measurements

Fig. 17 shows the uniform energy dissipation per-cycle (EPC) for the QRC\textsubscript{foot} across a range of operating frequencies with $V_{dd}$ fixed at $0.8$ V. Conventional clocking dissipation was conservatively estimated by downsizing the programmable clock driver to allow EPC measurements at iso-slew. This comparison also ignores any electromigration challenges associated with conventionally driving the entire clock load through a central buffer, and ignores the clock power dissipation that would be incurred in distributed clock buffers. The resulting energy efficiency-achieved QRC\textsubscript{foot} energy reduction compared to conventional clocking is also shown in Fig. 17. Measurements at $0.8$ V indicate EPC...
savings of 32%–39% compared to the traditional clocking. Since gate overdrive was not employed for \( M_c \) in QRC\textsubscript{foot}, higher voltages yield improved energy efficiency due to the more significant reduction in conduction losses.

To demonstrate importance of precisely timed gate control of \( M_c \), a manual override was used to sweep the delay codes for the \( t_{\text{rise}} \) and \( t_{\text{fall}} \) delay chains (Fig. 6) across a range of values and the resulting EPC reduction was measured.

Fig. 18 demonstrates the impact of \( M_c \) timing on energy-efficiency. Low delay codes result in premature de-assertion of \( M_c \), while late delay codes result in late \( M_c \) disconnection, both resulting in excessive conduction losses.

The energy efficiency of QRC\textsubscript{foot} under DVFS is shown in Fig. 19. The core supply voltage was swept from 1.2 to 0.7 V, and the on-chip core-supply powered ring oscillator provided the accompanying frequency scaling. QRC\textsubscript{foot} achieves an energy reduction of 32%–47% (across 0.7–1.2 V). QRC\textsubscript{foot} does not fundamentally limit scaling \( V_{\text{dd}} \) to 0.7 V or set a lower bound on the operating frequency in any way, a fact verified through simulation. The test-chip lower bounds in frequency and voltage in this test chip were the result of a design oversight pertaining to the use of level converters.

Fig. 20 shows measured EPC versus frequency at \( V_{\text{dd}} = 0.8 \) V. Fig. 21 shows EPC dissipation of QRC\textsubscript{pass} across a \( V_{\text{dd}} \) range of 0.7–1.2 V. Also included for reference are the conventional EPC numbers at corresponding voltages. Compared to traditional resonant clocking, QRC\textsubscript{pass} efficiency is relatively insensitive to frequency. Consistent with simulations, measured EPC values are lower than conventional clocking, but higher than QRC\textsubscript{foot}. Increased QRC\textsubscript{pass} EPC over QRC\textsubscript{foot} is largely attributable to the additional switching losses in the boosting circuit. Notably, QRC\textsubscript{pass} efficiency remains...
approximately constant across voltage and frequency scaling, a trend enabled by gate-boosting $M_G$.

Fig. 22 compares key design metrics of both QRC variants with related works. Overall, QRC demonstrates scaling across a wide voltage–frequency range. Near-uniform efficiency is achieved over a broad range of frequencies. Another important capability of QRC is its ability to achieve a broad duty-cycle range and ability to service multiple clock domains using a single inductor.

VIII. CONCLUSION

We presented the QRC architecture and demonstrated the first-ever continuous voltage–frequency scalable resonant clocking system. Autonomous timing control of QRC circuits plays a central role in enabling robust, efficient QRC operation. We presented test-chip implementation of two variants of QRC, implemented in 65-nm CMOS. Energy-efficiency measurements from the test-chip validate ability to QRC to achieve uniform efficiency across frequency scaling. DVFS measurements in the 0.7–1.2 V range indicate EPC reduction of 32%–47%, and 34%–38% using QRC$_{foot}$ and QRC$_{pass}$, respectively. Ability to achieve 0-cycle latency duty-cycle control offers a promising opportunity for QRC to be employed in a broader range of applications driving large capacitive loads beyond system clocking.

ACKNOWLEDGMENT

The authors would like to thank S. Pant, C. Tokunaga, S. Das, J. Koo, A. Smith, A. Loke, M. Khbeis, and J. Nasser for their valuable input during the course of this paper.

REFERENCES


Fahim ur Rahman (S’17) received the B.Sc. degree in electrical engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2013. He is currently pursuing the Ph.D. degree with the University of Washington, Seattle, WA, USA.

He was a Lecturer from 2013 to 2014. In 2014, he joined the University of Washington. His current research interest includes ultra-low power circuit design, advanced clocking techniques, phase-locked loop, MDLL, dc-to-dc power converters, and clock-power management for system clocking.

Viveesh Sathe (M’07) received the B.Tech. degree in electrical engineering from IIT Bombay, Mumbai, India, in 2001, and the M.S and Ph.D. degrees in electrical engineering and computer science from the University of Michigan, Ann Arbor, MI, USA, in 2004 and 2007, respectively.

From 2007 to 2013, he served as a Member of Technical Staff in the Low-Power Advanced Development Group at AMD, where his research focused on inventing and developing new technologies for the next-generation microprocessors. In 2013, he joined the Department of Electrical Engineering, University of Washington, Seattle, WA, USA.

Dr. Sathe has led the research and development effort at AMD that resulted in the first-ever resonant clocked commercial microprocessor. In addition, several of his other inventions have been adopted for use in future-generation microprocessors. His doctoral thesis was selected as the best dissertation in EECS for 2007 and was nominated for the Rackham Graduate School Distinguished Dissertation Award at the University of Michigan. He is a member of the Technical Program Committees of the Custom Integrated Circuits Conference and the International Symposium on Quality Electronic Design.