Behzad Razavi

# The Design of a Millimeter-Wave Frequency Divider

Following the design of a millimeter-wave VCO in the previous issue [1], we now turn to the feedback frequency divider that would be driven by the VCO in a PLL. We assume that the PLL receives a reference frequency of 50 MHz and generates an output frequency ranging from 28 to 32 GHz. Our target performance is as follows:

- *maximum input frequency*: 35 GHz
- *divide ratio*: 560–640
- *input voltage swing*: rail-to-rail and sinusoidal
- *power consumption*: <5 mW.

These specifications merit two remarks. First, even though the PLL's output nominally does not exceed 32 GHz, during settling it may—hence, the higher target of 35 GHz to ensure that the divider does not fail. Second, the divider input is assumed to be a sinusoid, a good approximation of the waveform generated by the VCO or its output buffer(s) at these frequencies.

We design the divider in 28-nm CMOS technology with a worst-case supply of $1\ V - 5\% = 0.95\ V$ at 75 °C and in the slow–slow corner. All (drawn) transistor lengths are chosen equal to 30 nm.

## Multimodulus Dividers

In this design, we pursue the "modular" architecture introduced in [2]. To arrive at this topology, let us begin with the ÷3 circuit shown in Figure 1(a). Two DFFs and an AND gate form the loop and produce $Q_1\,\overline{Q}_2 = 01, 11, 10$ as the clock cycles arrive. In this article, FFs and latches are represented by double and single boxes, respectively.

Next, we modify the topology so that it can divide by two or three according to a "modulus control" (MC) input. As depicted in Figure 1(b), we return the output from $Q_2$ to the first AND gate and compensate for this negation by sensing $\overline{Q}_1$. This ÷2/3 stage (Div23) divides the clock frequency by three if $MC = 1$ and by two if $MC = 0$, the latter because $FF_2$ and the AND gate then act as a single inverting loop, while $\overline{Q}_1 = 1$.

We now cascade a Div23 stage with a ÷2 circuit and return the latter's inverted output to the former's MC input [Figure 2(a)]. Assuming that the stages change their outputs on the rising edges of their respective inputs, suppose we begin with $\overline{Q}_2\,\overline{Q}_3 = 01$. When *CK* goes

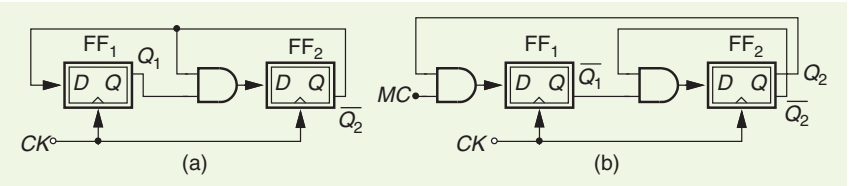


**FIGURE 1:** The (a) divide-by-three and (b) dual-modulus circuits. MC: modulus control.

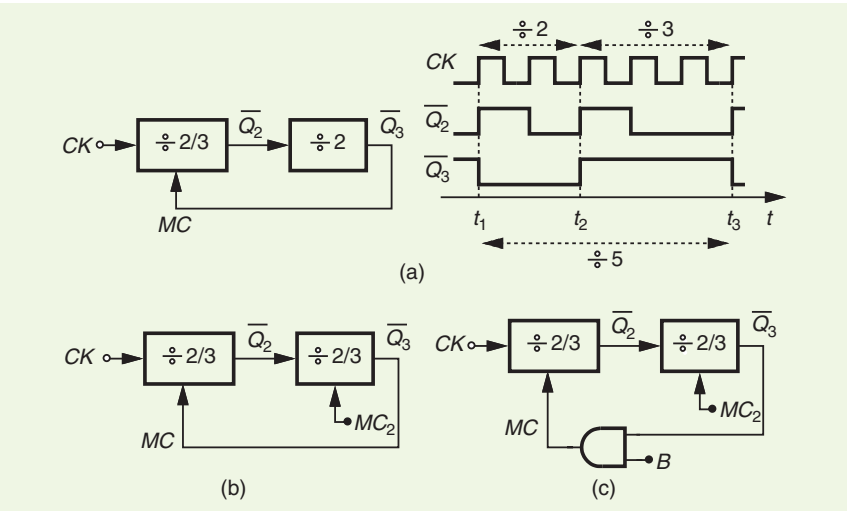


**FIGURE 2:** (a) A ÷2/3 and a ÷2 block in a loop. (b) Two ÷2/3 blocks in a loop, and (c) the circuit of (b) with an additional control input.

*Digital Object Identifier 10.1109/MSSC.2022.3205805*
*Date of current version: 11 November 2022*

high at $t = t_1$, so does $\bar{Q}_2$, and this rise causes $\bar{Q}_3 = MC$ to change as well. Div23 then divides by two, bringing $\overline{Q}_2$ back to a high level at $t = t_2$. Now, $\overline{Q}_3$ changes, and Div23 divides by three until $t = t_3$. We observe that the overall circuit acts as a $\div 5$ stage. One can readily show that substituting the $\div 2$ block with a $\div 3$ structure creates a divide ratio of seven. Thus, the topology illustrated in Figure 2(b) divides by five if $MC_2 = 0$ and by seven if $MC_2 = 1$.

Depicted in Figure 2(c) is a more general case where the divide ratio is equal to four if $B = MC_2 = 0$ (because both stages reduce to $\div 2$ blocks) and equal to six if $B = 0$ and $MC_2 = 1$ (because the second stage divides by three). In summary, this arrangement can divide by four, five, six, or seven according to the logical levels applied to $B$ and $MC_2$.

## Divider Architecture

To develop a modular chain, we redraw the $\div 2/3$ circuit of Figure 1(b) at the latch level in Figure 3(a) but with one modification. We insert an AND gate between $L_1$ and $L_1'$ and control it by an input $B$. The circuit divides by three if $MC = B = 1$ and by two if either is low.

In the last step of our development, we employ two instances of this topology in the form suggested by Figure 2(c). Illustrated in Figure 3(b), the result receives a "static" digital input $B_1 B_2$ that defines its divide ratio. Also, the MC of the first module is driven by latch $L_1$ in the following stage [as in Figure 2(b)]. It can be shown that for $n$ such stages, the divide ratio is equal to $2^n + 2^{n-1} B_n + \cdots + 2^0 B_1$, where $MC_n$ is assumed to be high [2]. This value can go from $2^n$ to $2^{n+1} - 1$ in steps of unity. For our target range of 560–640, we must select $n = 9$.

## Choice of Logic Style

With a maximum input frequency of 35 GHz, we must implement the divider with a logic style that robustly supports the speed and yet consumes low power. We prefer to avoid CML for its static power and surmise that CMOS (rail-to-rail) logic can deliver the required performance.

We examine two CMOS logic styles here, namely, two versions of the clocked CMOS (C²MOS) structure [3]. To quantify their speed limitations, we use these styles to construct $\div 2$ circuits. Figure 4 depicts the two realizations along with their transistor dimensions and an output buffer. Note that the $C^2MOS_A$ and $C^2MOS_B$ versions simply differ by how they accommodate the clocked devices in the stack.

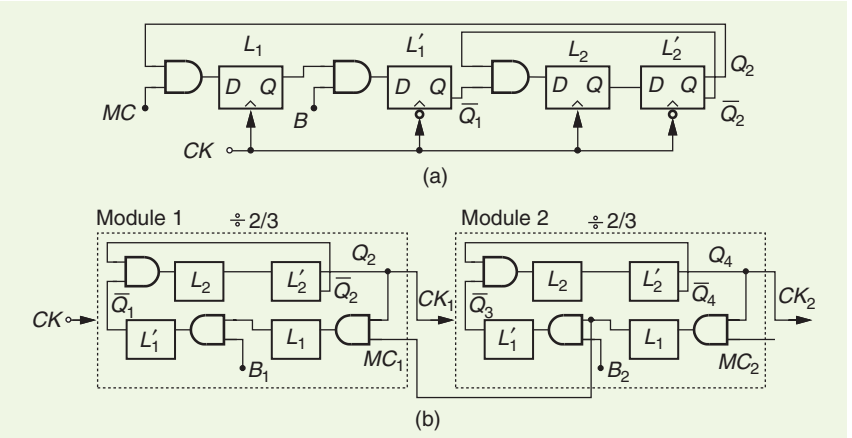


**FIGURE 3:** (a) The ÷2/3 circuit at the latch level and (b) two such stages in a loop.

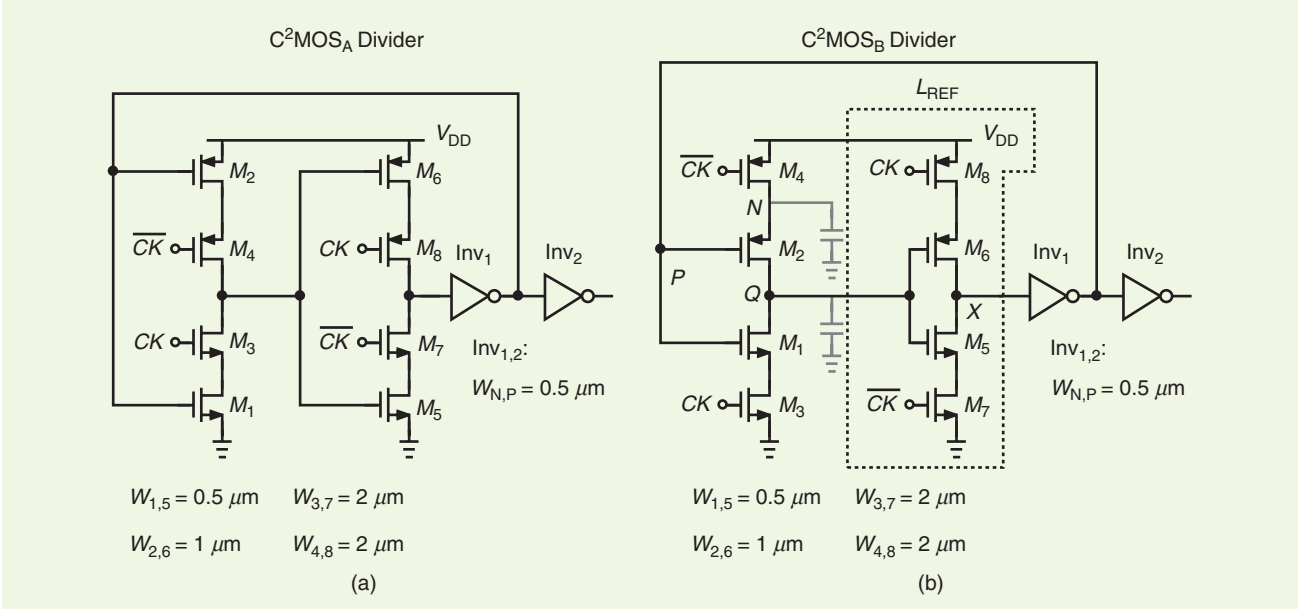


**FIGURE 4:** The C²MOS latches with clocked devices (a) in the signal path and (b) outside the signal path.

Before evaluating these circuits' maximum toggle frequencies, we should make three remarks. First, in Figure 4(a), the clocked transistors introduce capacitance directly in the signal path, whereas in Figure 4(b), they do not. We then expect $C^2MOS_B$ to achieve a higher speed. Second, this structure, unfortunately, suffers from charge sharing. Suppose both *CK* and *Q* are low while *P* goes low. Transistor $M_2$ then turns on, allowing the charge on the capacitance at node *N* to be shared with that at *Q*. The low level at this node, therefore, degrades significantly. A similar effect occurs if *Q* is high and *P* rises. Figure 5 plots the voltage waveforms at *Q* and *N* for an input frequency $f_{in}$ of 30 GHz, revealing the kinks that arise from charge sharing. At very high speeds, nevertheless, the kinks begin to disappear.

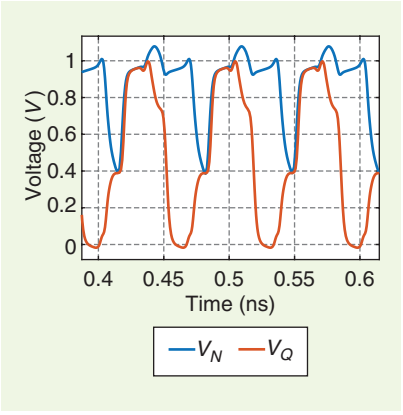


**FIGURE 5:** The divider output showing kinks in transitions.

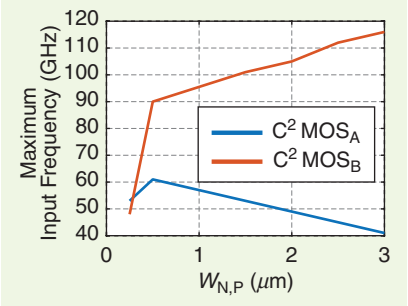


**FIGURE 6:** The maximum divider speed versus the width of clocked transistors.

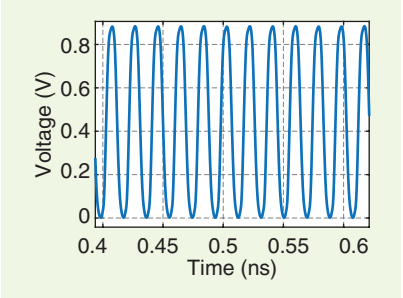


**FIGURE 7:** The type B divider output with $f_{in}$ = 105 GHz.

Third, in optimizing frequency dividers, we typically observe a higher performance as we increase the widths of the clocked transistors—because they present less resistance—but we must also consider the power consumption in the clock path. Expressed as $f_{in} C_{in} V_{DD}^2$, this value can, in fact, exceed that burnt by the data path. A fair comparison of different topologies, therefore, requires the same total $C_{in}$ for each. In Figure 4, the total transistor width in the clock path is equal to 8 μm.

We examine the maximum speed of the $C^2MOS_A$ and $C^2MOS_B$ ÷2 circuits as the width of the clocked transistors is varied. Simulations yield the behavior illustrated in Figure 6. We observe the following:

- The A style exhibits an optimum; this is because wider transistors, in this case, also introduce greater capacitance in the signal path.
- For the B style, the speed improves up to the point where the on-resistance of the clocked transistors is several times less than that of the devices in the data path.
- Except for the first point on the left, the B version remains faster than its A counterpart, confirming our intuition that placing the clocked devices directly in the data path degrades the speed.
- For a width of 2 μm, the B structure runs up to an input frequency of 105 GHz.

We then select the B version for the first stage in Figure 3(b) with a width of 2 μm while bearing in mind its charge-sharing issue. Figure 7 depicts the waveform at node *X* in Figure 4(b) with $f_{in}$ = 105 GHz. We shall call $L_{REF}$ in this figure the "reference" latch, as it provides the optimal ratios of transistor widths.

We should raise three other points here. First, the 61-GHz limit facing the A version appears to lie well above the range of interest to us, suggesting that this type should suffice for our purpose. However, layout parasitics drastically reduce the speed. Moreover, ÷2 arrangements make optimistic predictions for circuits having greater divide ratios. For these reasons, we resort to the B style for the first module and design the entire divider for $f_{in}$ = 60 GHz, expecting that it can operate at 35 GHz after it is laid out.

Second, in contrast to the conceptual latches in Figure 3(a), $C^2MOS$ topologies do not provide complementary outputs. As seen in the following sections, the module design must be modified accordingly.

Third, given that the second module in Figure 3(b) presents a large capacitance to the output of the first and that it operates at a lower speed, we wonder whether it can employ narrower clocked devices. We return to this point later in the article.

## Design of the First Module

The first module in Figure 3(b) poses the greatest speed challenge. To accommodate NAND functions, we modify the topology as illustrated in Figure 8(a). Inverters $Inv_1$ and $Inv_2$ do introduce additional delays, but the alternative would be to employ NOR gates, which are substantially slower. Also, these inverters' small transistors require two-stage buffers $Inv_{3,5}$ and $Inv_{4,6}$ for driving the input capacitance of the next module.

Each NAND/latch combination is realized as shown in Figure 8(b). Note that the transistor widths in the data path are scaled with respect to those in the reference latch, $L_{REF}$ in Figure 4(b). The single latch, $L_2'$ in Figure 8(a), is based on $L_{REF}$ as well.

We first perform simulations to determine the maximum speed of this module. We surmise that the worst case occurs for ÷3 operation and examine the waveforms at nodes *E* and *F* in Figure 8(a). Plotted in Figure 9(a) for $f_{in}$ = 60 GHz, these waveforms reveal proper operation. The kinks observed at *F* are removed by the following inverters. Figures 9(b) and (c) show the results in the ÷2 and ÷7

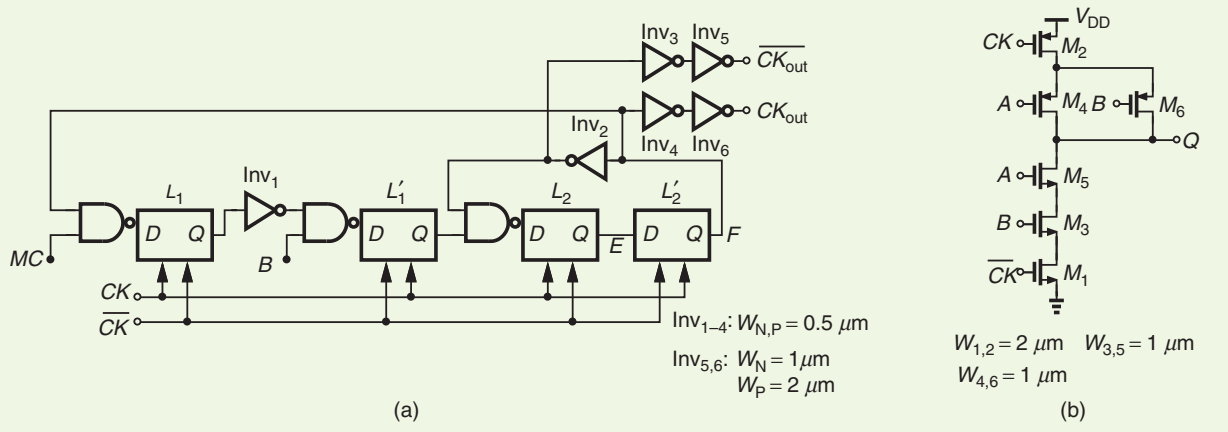


**FIGURE 8:** (a) The latch-level implementation of module 1 and (b) the transistor-level design of the NAND/latch circuit.

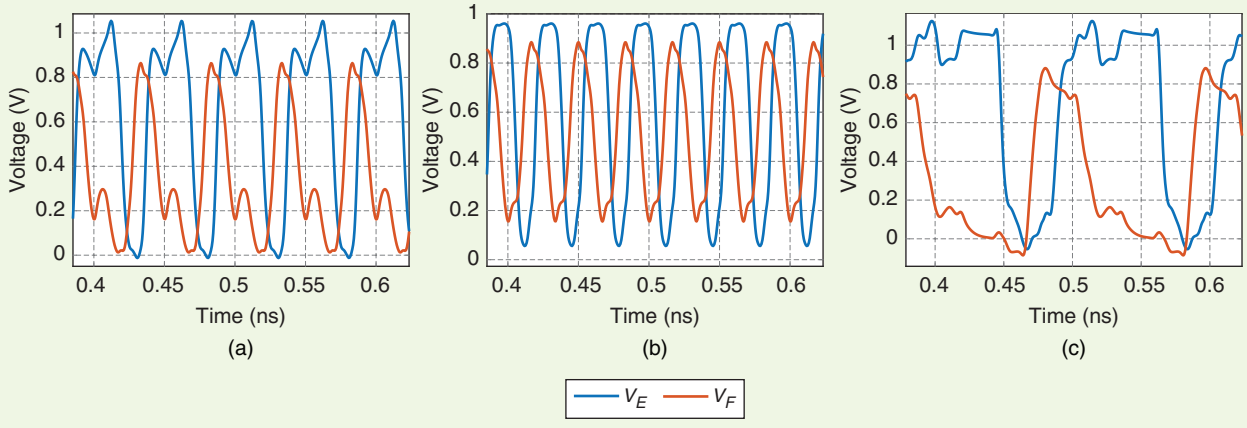


**FIGURE 9:** The first module output waveforms in (a) ÷3, (b) ÷2, and (c) ÷7 modes for $f_{in}$ = 60 GHz.

modes, respectively. Charge-sharing phenomena are not pronounced at these frequencies.

We now apply a 30-GHz input and study the effect of charge sharing. Depicted in Figure 10(a), the waveforms at *E* and *F* exhibit some kinks. However, we wonder whether $Inv_2$ and $Inv_4$ in Figure 8(a) incur greater phase noise (PN) due to the kinks at their inputs. To answer this question, we study the PN at the output of $Inv_2$ in two different cases: (a) the module operates as a ÷3 circuit or (b) the input of this inverter is driven by an ideal voltage source having smooth transitions and rise and fall times equal to 10 ps. As Figure 10(b) indicates, the PN penalty arising from the kinks amounts to about 4 dB. Nonetheless, this PN is still far below that obtained for the millimeter-wave VCO designed previously. For a fair comparison, we raise the degraded PN profile in Figure 10(b) by $20\log 3 = 10$ dB so as to refer it to the divider input and note that the result is around –135 dBc/Hz at 1-MHz offset. The VCO PN, on the other hand, is approximately equal to –110 dBc/Hz at this offset.

Having reached a satisfactory first-stage design, we now compute its power consumption. According to simulations, the module of Figure 8(a) draws 0.4 mW at $f_{in}$ = 30 GHz in the

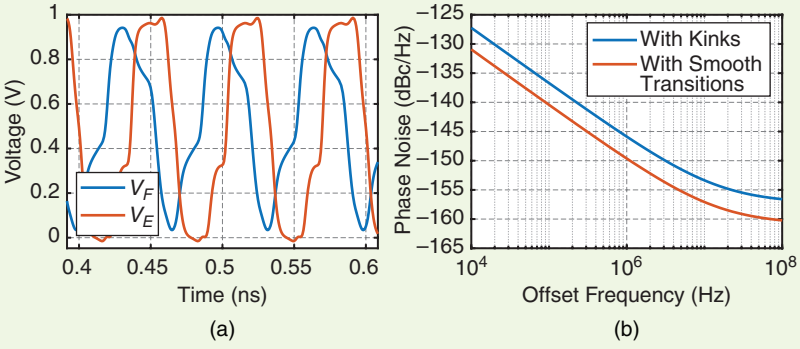


**FIGURE 10:** (a) The internal waveforms of module 1 for $f_{in}$ = 30 GHz and (b) the phase noise profile for the actual circuit and for the case of ideal transitions at node *F*.

÷2 mode. However, we must also deal with the clock path, which faces a total transistor width of 16 $\mu$m.

Two clock buffer alternatives emerge at this point. First, we can incorporate an *LC*-tuned stage to absorb the divider input capacitance. This solution consumes little power but leads to a more complex layout. Second, we can simply insert relatively large inverters in the *CK* and $\overline{CK}$ paths, thus incurring $P = f_{in} C_{in} V_{DD}^2$. Shown in Figure 11, the inverter dimensions are so chosen as to guarantee rail-to-rail swings at the divider inputs. Simulations reveal a total power of 0.92 mW for the two inverters at $f_{in} = 30$ GHz, an amount greater than that of the core. We expect that layout parasitics nearly double this amount.

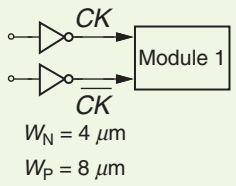


**FIGURE 11:** The input buffers for evaluating the clock path power consumption.

## Addition of the Second Module

As Figure 3(b) suggests, we can simply follow the first module by a copy of itself, with $CK_{out}$ and $\overline{CK_{out}}$ in Figure 8(a) driving the second module [Figure 12(a)]. However, the lower speeds involved here make charge sharing and the resulting kinks more serious. For this reason, we implement the second module using the C$^2$MOS$_A$ style. Figure 12(b) depicts the building blocks. Chosen according to the optimum suggested by Figure 6, the clocked transistor widths are smaller than those in the first module, thereby reducing the capacitance seen by $CK_{out}$ and $\overline{CK_{out}}$.

It is important to note that the connection of $CK_{out}$ and $\overline{CK_{out}}$ between the two modules is not arbitrary. In other words, if these two outputs are swapped, the chain fails. To appreciate this point, suppose we have $B_1 = 1$ and $B_2 = 0$ in Figure 12(a) so that the cascade divides the main input frequency by five. Recall from Figure 2(a) that both stages change their outputs on the rising edge of their respective inputs. That is, $\overline{Q_3}$ must change only when *CK* and $\overline{Q_2}$ go high. In the realization of Figure 12(a), this condition is fulfilled by applying $CK_{out}$ to the *CK* input of the first and third latches in the second module.

The second module operates at a lower speed, thus consuming 0.2 mW. We can further reduce the power by redesigning the second module with smaller transistors, but the savings are negligible.

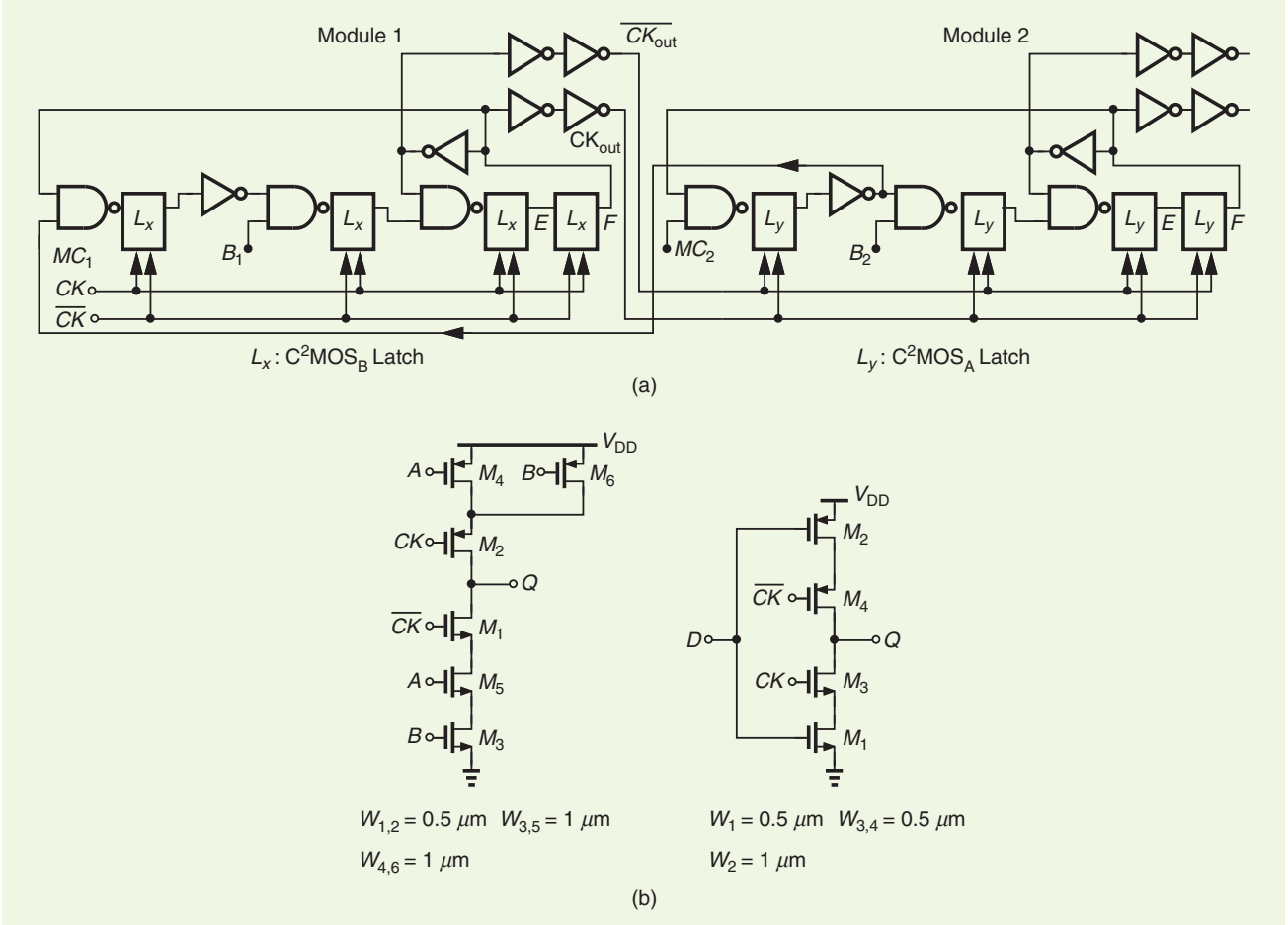


**FIGURE 12:** (a) The cascade of the first two modules and (b) the design of the second module's building blocks.

*(continued on p. 16)*

4) Large filtering caps will destabilize your feedback loop.

Want to rethink your career choice?

## Take a Deep Breath

It's standard editorial policy, even in mainstream news magazines, to capture the reader's attention with frightening headlines or alarming stories of gloom and doom but then to offer some potential solution or some small ray of hope so readers leave the story with a grateful sense of relief that maybe things will turn out all right after all. Not here. It's good for engineers to be anxious. Mother Nature and her evil brother Murphy *are* out to get you. You need to keep your guard up.

There is no easy fix for VDD problems. While we often get by with just dumb luck, these problems will not fix themselves. And chances are, they will eventually trip you up... if they haven't already.

That's not to say there is nothing you can do. For example

- Come up to speed on PDN issues. Reference [2] is an excellent source of practical info and is quite easy to read. Large SoC teams will usually have somebody assigned to worry about this stuff. Make contact. There may be a supply model you can simulate with. Don't expect such models to be too precise, though. Resonant frequencies can only be estimated, even when the layout is known.
- Check the layout. Do the VDD or GND planes look "fingery" and inductive?
- People working at high speed are likely to take steps to damp their on-chip bypass caps. See [5] for an efficient solution. You may also want to take a look at "The 250-ps Speed Limit," because capacitors are not always what you expect.
- Good circuit practices can eliminate—or at least mitigate—many troublesome sources of supply coupling. If you'd like to see an article on this topic in this column, send your requests here: shifobrains@ieee.org. As always, be sure to include flattering comments about how much you love this series.
- Simulate PSRR, but remember that simulators think everything is perfectly balanced. Real circuits are not. Imbalance your differential or pseudo-differential circuit with mismatch and signal swing.
- Understand how your feedback circuit can become inductive. See [4] for more on this topic.

OK, that's it for now. I leave you with the parting wisdom of Elvira, Mistress of the Dark whose tagline was "Unpleasant dreams!"

## References


[1] G. Kim, D. Lim, T. Das, E. Lee, and S. You, "System Level Power Supply Induced Jitter Suppression for multi-lane high speed serial links," in *Proc. IEEE 72nd Electron. Compon. Technol. Conf. (ECTC)*, 2022, pp. 1863–1868, doi: 10.1109/ECTC51906.2022.00293.

[2] L. D. Smith and E. Bogatin, *Principles of Power Integrity for PDN Design – Simplified*. Englewood Cliffs, NJ, USA: Prentice Hall, 2017.

[3] "Catch-22." Dictionary. [Online]. Available: https://www.dictionary.com/browse/catch-22

[4] C. Mangelsdorf, "Oscillation and ringing: Finding worst-case capacitive load [Shop Talk: What You Didn't Learn in School]," *IEEE Solid-State Circuits Mag.*, vol. 13, no. 2, pp. 16–20, Spring 2021, doi: 10.1109/MSSC.2021.3072801.

[5] C. Mangelsdorf, "Duct Tape for the Frequency Domain: If This Ol' Hack Don't Fix It, It Ain't Broke [Shop Talk: What You Didn't Learn in School]," *IEEE Solid-State Circuits Mag.*, vol. 13, no. 4, pp. 15–22, Fall 2021, doi: 10.1109/MSSC.2021.3111427.


## THE ANALOG MIND *(continued from p. 10)*

### Complete Divider

We are now ready to create a chain of nine modules so as to meet the necessary divide ratio range. Note that the MC input of the ninth module must be tied to $V_{DD}$. The overall circuit draws a power less than 3 mW. Figure 13 plots the waveforms at nodes $E$ and $F$ of this module with $f_{in}$ = 30 GHz and for an overall divide ratio of 575. Let us examine the droop seen in these voltages.

The dynamic nature of C²MOS logic makes it susceptible to leakage currents drawn by the transistors, an issue that becomes more serious for long clock periods. It also manifests itself to a greater extent at high temperatures, as the subthreshold leakage rises exponentially. In our design, the last module must generate an output frequency equal to 50 MHz if the divider operates within a PLL sensing a 50-MHz reference. With the C²MOS latches residing in the store mode for 10 ns, their output states may degrade significantly. In the waveforms of Figure 13, the waveform at node $E$ suffers from a droop of about 300 mV. Nonetheless, the inverters following this node restore the logical levels.

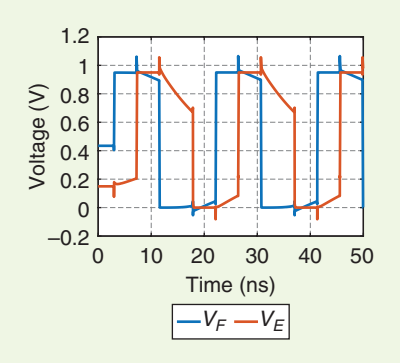


**FIGURE 13:** The output waveforms of the complete divider exhibiting the effect of leakage.

## References


[1] B. Razavi, "The design of a millimeter-wave VCO [The Analog Mind]," *IEEE Solid-State Circuits Mag.*, vol. 14, no. 3, pp. 6–12, Summer 2022, doi: 10.1109/MSSC.2022.3184443.

[2] C. S. Vaucher et al., "A family of low-power truly modular programmable dividers in standard 0.35-µm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 35, no. 7, pp. 1039–1045, Jul. 2000, doi: 10.1109/4.848214.

[3] Y. Suzuki, K. Odagawa, and T. Abe, "Clocked CMOS calculator circuitry," *IEEE J. Solid-State Circuits*, vol. 8, no. 6, pp. 462–469, Dec. 1973, doi: 10.1109/JSSC.1973.1050440.