Distributed Integrated Circuits: An Alternative Approach to High-Frequency Design

Ali Hajimiri, California Institute of Technology, Pasadena

ABSTRACT

Distributed integrated circuits are presented as a methodology to design high-frequency communication building blocks. Distributed circuits operate based on multiple parallel signal paths working in synchronization that can be used to enhance the frequency of operation, combine power, and enhance the robustness of the design. These multiple signal paths usually result in strong couplings inside the circuit that necessitate a treatment spanning architecture, circuits, devices, and electromagnetic levels of abstraction.

INTRODUCTION

"Divide and conquer" has been the underlying principle used to solve many engineering and social problems. Over many years engineers have devised systematic ways to divide a design objective into a collection of smaller projects and tasks defined at multiple levels of abstraction. This approach has been quite successful in an environment where a large number of people with different types and levels of expertise work together to realize a given objective in a limited time. Communication system design is a perfect example of this process, where the communication system is initially defined at the application level, then described using system level terms, leading to an architecture using a number of cascaded sub-blocks that can be implemented as integrated circuits. The integrated circuit design process is then divided further by defining the specifications for circuit building blocks and their interfaces that together form the system. The circuit designer works with the specifications at a lower level of abstraction dealing with transistors and passive components whose models have been extracted from the measurements, device simulations, or analytical calculations based on the underlying physical principles of semiconductor physics and electrodynamics. This process of breaking down the ultimate objective into smaller, more manageable projects and tasks has resulted in an increase in the number of experts with more depth yet in more limited sublevels of abstraction.

While this divide-and-conquer process has been quite successful in streamlining innovation, the overspecialization and short time spans associated with today's design cycles sometimes result in suboptimal designs in the grand scheme of things. Also, in any reasonably mature field many of the possible innovations leading to useful new solutions within a given level of abstraction have already been explored. Further advancements beyond these local optima can be achieved by looking at the problem across multiple levels of abstraction to find solutions not easily seen when one confines one's search space to one level (e.g., transistor-level circuit design). This explains why most of today's research activities occur at the boundaries between different levels of abstraction artificially created to render the problem more tractable.

Distributed circuit design is a multilevel approach allowing a more integral co-design of the building blocks at the circuit and device level. Unlike most conventional circuits, it relies on multiple parallel signal paths operating in harmony to achieve the design objective. As will be shown in this article, this approach offers attractive solutions to some of the more challenging problems in high-speed communication circuit design. We briefly describe some of the issues in communication transceivers, wired or wireless, and then discuss some of the underlying features of distributed circuits and the associated design methodology. Finally, we demonstrate these concepts through three different examples of practical distributed building blocks.

0163-6804/02/$17.00 © 2002 IEEE

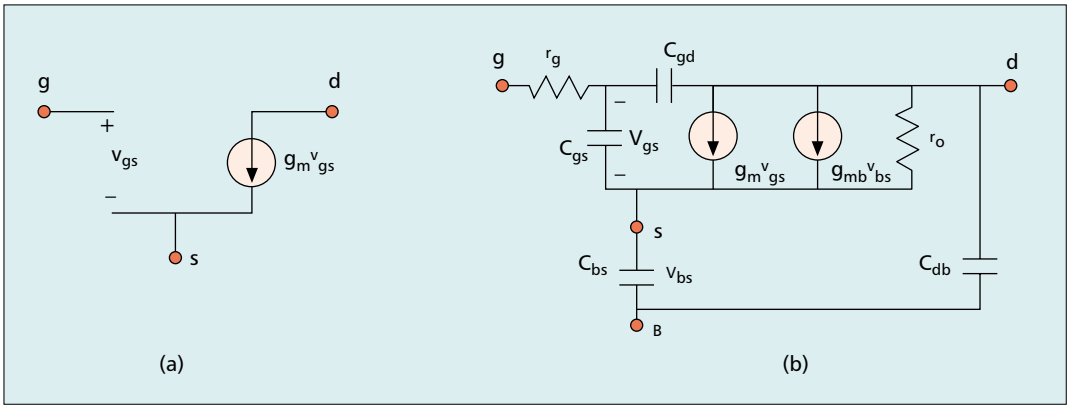

Figure 1. *a) The basic small-signal model of an MOS transistor; b) the model enhanced by addition of various parasitic components.*

Conventional lumped circuit design is a very well defined level of abstraction between device physics and system theory. In conventional electronic design, circuit elements such as transistors, resistors, and capacitors are defined by their voltage-current characteristics in time and frequency domain.

Issues in High-Speed Integrated Communication Circuits

Integration of high-speed circuits for wireless (e.g., cellular phones) and wired applications (e.g., optical fiber communications) poses several challenges. High-speed analog integrated circuits used in wireless and wired communication systems have to achieve tight and usually contradictory specifications. Some of the most common specifications are the frequency of operation, power dissipation, dynamic range, and gain. Once in a manufacturing setting, additional issues, such as cost, reliability, and repeatability, also come into play. To meet these specifications, the designer usually has to deal with physical and topological limitations caused by noise, device nonlinearity, small power supply, and energy loss in the components.

Frequency of operation is perhaps one of the most important properties of communication integrated circuits since a higher frequency of operation is one of the more evident methods of achieving larger bandwidth, and hence higher bit rates in digital communication systems. A transistor in a given process technology is usually characterized by its unity-gain frequency shown as f_T. This is the frequency at which the current gain of a transistor drops to unity. While the unity-gain frequency of a transistor provides an approximate measure to compare transistors in different process technologies, the circuits built using these transistors scarcely operate close to the f_T and usually operate at frequencies 3–100 times smaller depending on the complexity of their function. There are two main reasons for this behavior. First, analog building blocks and systems usually rely on closed loop operation based on negative feedback to perform a given function independent of the parameter variations. An open loop gain much higher than one is thus required for the negative feedback to be effective. Even if no feedback is present and open loop operation is possible, higher gain usually improves the noise and power efficiency of the circuits. Therefore, the transistor has to operate at a frequency lower than the f_T to provide the desired gain. Second, passive devices (e.g., capacitors and inductors), necessary in most high-speed analog circuits, have their own frequency limitations due to parasitic components that can become the bottleneck of the design. The combination of these two effects significantly lowers the maximum frequency of reliable operation in most conventional circuit building blocks and provides a motivation to pursue alternative approaches to alleviate the bandwidth limitations.

Low power operation is another desired feature of communication systems, particularly in portable wireless applications where the power dissipation directly determines the battery life. Even in wired systems (e.g., optical communications) where more power is usually available, the thermal control and heat dissipation limit the number of transceivers that can operate in close proximity to one another (e.g., fiber bundles). The power dissipation depends on a large number of parameters and is generally a complex function of its specifications. Distributed circuit design can be used to overcome the frequency and power limitations of conventional circuits.

Conventional Circuit Design

Conventional lumped circuit design is a very well defined level of abstraction between device physics and system theory. In conventional electronic design, circuit elements such as transistors, resistors, and capacitors are defined by their voltage-current characteristics in time and frequency domain. These characteristics are nonlinear in general. The accuracy of the analysis can be improved by including the parasitic components caused by second order effects in the device model. As an example, Fig. 1 shows the simplest small-signal model for a metal oxide semiconductor (MOS) transistor and an improved version including some parasitic components.

As long as the circuit dimensions are much smaller than the wavelengths involved, the circuit is called lumped and the entire network of elements can be modeled completely and accurately using the Kirchhoff's voltage and current laws (KVL and KCL) combined with the device voltage-current characteristics. The resultant sets of equations can be solved numerically and/or analytically. This enables the circuit designer to think about components such as transistors, resistors, and capacitors as standalone building blocks and form intuition about their behavior at

The term distributed has been used in various contexts with different meanings. In our treatment, a system will be called distributed if it uses multiple parallel signal paths and devices working in harmony to perform a desired task.

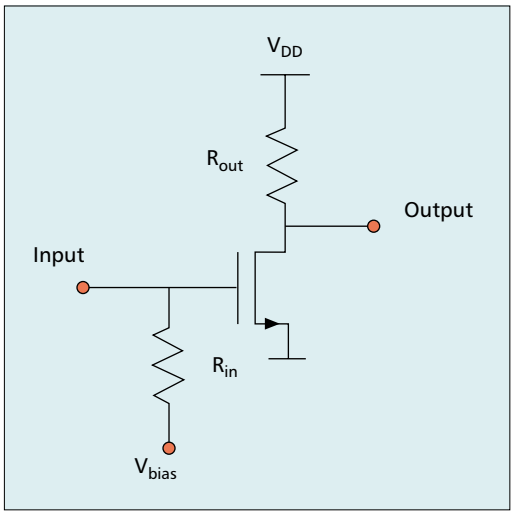

Figure 2. *A conventional common-source amplifier.*

the circuit level, without worrying about the details of device physics or higher-level system properties involved (e.g., the electron transport in a transistor or the hand-off process in a wireless network, respectively). This approach has led to substantial progress in the field of electronic circuit design resulting in large-scale analog and mixed signal integrated systems. Also, a variety of computer aided design (CAD) tools have been developed for the purpose of simulation, layout, and verification of such large chips.

Another interesting feature of most conventional analog integrated systems and circuits is their cascaded (series) nature. The system function is often achieved by a chain of blocks operating on the signal in series, where the output of each block is the input to the next one. In most systems there is a *single* (or a *pair* of complementary) signal-bearing path(s) through which the information is processed. This cascaded structure of transceivers is definitely in the spirit of the divide-and-conquer mentality.

As an example, Fig. 2 shows a conventional common source amplifier providing voltage and power gain from its input to its output. The gain of this amplifier is determined by the transistor's transconductance, g_m, and the resistances. Its bandwidth is determined by the RC time constants associated with the resistors and the parasitic capacitors of the transistor shown in Fig. 1b. This topology has a very strict trade-off between its gain and bandwidth, where improving one proportionally degrades the other. The gain-bandwidth product of the amplifier is usually a measure of the speed of the transistor and the effectiveness of the amplifier topology.

Distributed Circuit Design

Distributed circuits can defy some of the performance trade-offs in conventional circuits by taking advantage of multiple parallel signal paths. This multiple signal path feature of the distributed systems often results in strong electromagnetic couplings between circuit components across multiple levels of abstraction. This strong coupling makes it necessary to perform the analysis and design of distributed circuits across different levels. Due to the larger solution space, the design may be closer to a global optimum and hence outperform designs based on conventional circuit design. It is obvious that such multilevel approaches are more challenging due to the larger number of disciplines that need to be mastered, as well as a current lack of proper CAD tools.

The term *distributed* has been used in various contexts with different meanings. In our treatment, a system will be called distributed if it uses multiple parallel signal paths and devices working in harmony to perform a desired task. This is in contrast to the more conventional serial cascaded systems with a single (or pair of) signal path(s). Under this rather liberal definition of a distributed system, the actual physical dimensions may or may not be comparable to the wavelengths involved.[1] As seen in the rest of this article, distributed circuits can achieve higher operation frequency, lower sensitivity to passive energy losses, and more robustness to process variations.

This concept can be best seen through an example. Figure 3a shows the distributed amplifier originally suggested by Percival [2] and then by Ginzton *et al.* [3]. The operation of the amplifier can be understood more easily using the transmission-line-based version of the amplifier shown in Fig. 3b. This distributed amplifier consists of two transmission lines on the input and the output, and multiple transistors providing gain through multiple signal paths. The forward wave (to the right in the figure) on the input line is amplified by each transistor. The incident wave on the output line travels forward in synchronization with the traveling wave on the input line. Each transistor adds power in phase to the signal at each tap point on the output line. The forward traveling wave on the gate line and the backward (traveling to the left) wave on the drain line are absorbed by terminations matched to loaded characteristic impedance of the input line, R_{in}, and output line, R_{out}, respectively, to avoid reflections.

For input and output lines with equal characteristic impedances, the gain of the distributed amplifier can be approximated by [4]

$$A_v \approx n/2 \cdot g_m R \cdot L,$$

where n is the number of transistors, g_m is the transconductance of each transistor, R is the characteristic impedances of the input and output lines, and L is the end-to-end loss of the effective transmission lines. As can be seen, in the absence of loss the gain can be increased by addition of extra transistor sections without a penalty on the bandwidth and hence improving the gain-bandwidth product without a limit.

The transmission lines may be replaced with lumped inductors, as in Fig. 3a. These inductors together with the parasitic capacitors of the transistors form an LC ladder low-pass filters at the input and output. These ladder low-pass filters can be designed to provide a bandwidth in excess of that obtained with the lumped implementation shown in Fig. 2 [3].[2]

It is noteworthy that at low frequencies the distributed amplifiers of Fig. 3 are equivalent to the conventional common source amplifier of Fig. 2 with multiple parallel transistors. However, sim-

[1] *Although some authors consider distributed the opposite of lumped, the term distributed has been used to referred to circuits with multiple parallel signal paths from the very early days of distributed circuits (e.g., see [1]).*

[2] *The mechanical counterpart of the LC ladder filter is an array of equally spaced identical springs and weights. This problem was first analyzed by Newton where he shows that this system has a cutoff frequency above which the mechanical wave does not propagate in the normal fashion. Later, Johann Bernoulli and Lord Rayleigh expanded the study to the case where the spacing and size of the weights become progressively smaller, resulting in a continuous system similar to the transmission-line-based distributed amplifier of Fig. 3b.*

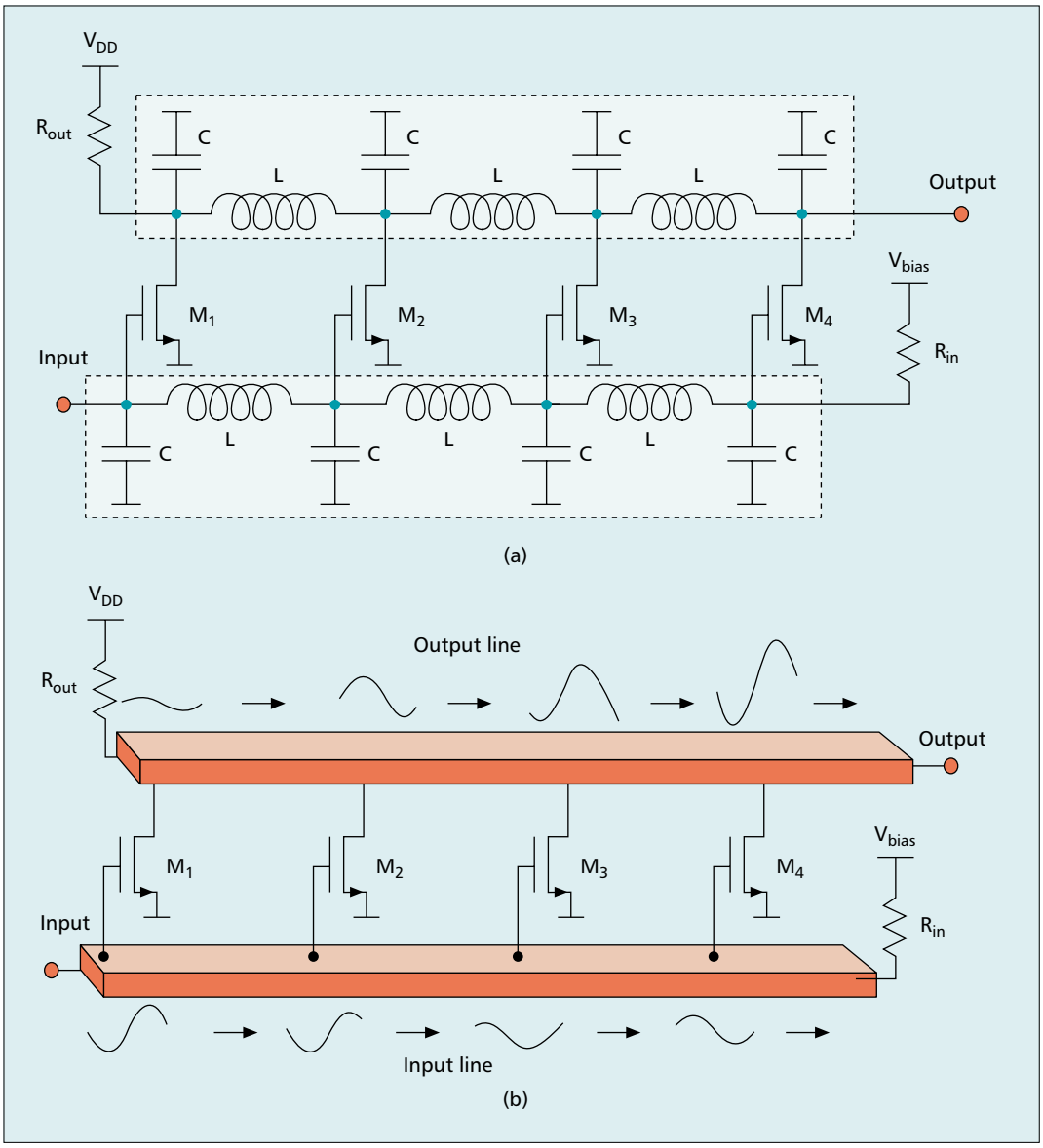

Figure 3. *A distributed amplifier with LC ladder filter input and output line; b) a distributed amplifier with transmission line sections.*

A distributed amplifier has a relaxed gain-bandwidth trade-off compared to a lumped amplifier since the parasitic capacitances of the transistor are absorbed into the transmission lines or the LC ladder filter to become a part of the passive network.

ply using *n* common source amplifiers in parallel does not change either the voltage gain or the bandwidth of the amplifier. This clearly demonstrates that it is the multiple signal path nature of the distributed systems that improves bandwidth.

A distributed amplifier has a more relaxed gain-bandwidth trade-off than a lumped amplifier since the parasitic capacitances of the transistor are absorbed into the transmission lines or the LC ladder filter to become part of the passive network. These parasitics are thus canceled to the first order by the interstage inductors or transmission lines. The distributed amplifier can provide larger-than-unity overall voltage gain even at frequencies where each transistor has sub-unity gain as the transistor contributions to the gain add directly on the output line. In the absence of passive loss, additional gain can be achieved without a significant reduction in bandwidth by addition of extra transistor segments. This is the direct result of multiple signal paths in the circuit. The extended bandwidth of the distributed amplifier comes at the price of a larger time delay between its input and output. It can be shown that there is a trade-off between the bandwidth and delay in an amplifier [5]. Distributed amplification provides a means to take advantage of this trade-off in applications where the delay is not a critical specification of the system and can be compromised in favor of the bandwidth.

It is noteworthy that the physical size of a distributed amplifier does not have to be comparable to the wavelength for it to enhance the bandwidth. In fact, the first demonstration of distributed amplification used purely lumped elements similar to Fig. 3b[3]. As discussed earlier, the highest frequency of operation for an amplifier is generally limited by the parasitics associated with the passive and active components. Dividing the gain between multiple active devices avoids the concentration of the parasitics at one place and hence eliminates a dominant pole scenario in the frequency domain transfer function.[3] Here, one tries to avoid a weakest link

[3] *This concept can be stated in very loose sense as "dividing the evil," which can be found in many engineering problems.*

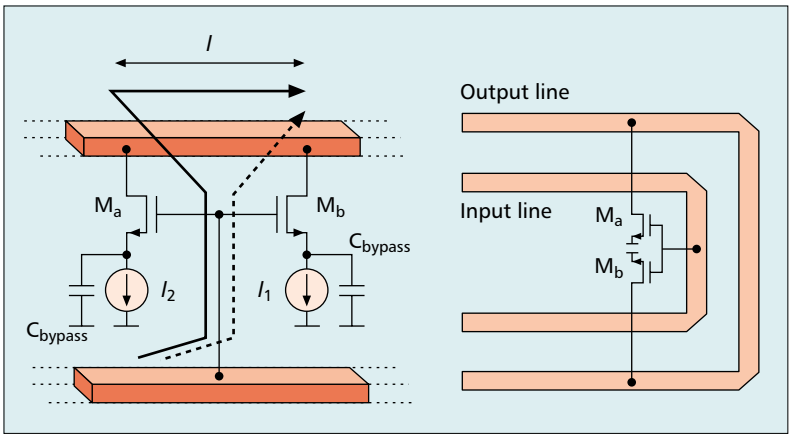

Figure 4. *One tuning section for frequency control of the distributed oscillator.*

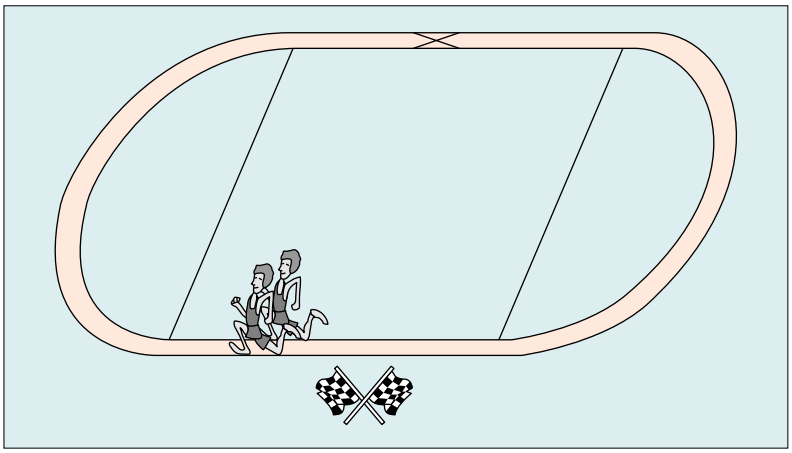

Figure 5. *The race track analogy for the tuning technique used in the DVCO.*

situation by providing multiple equally strong (or equally weak) parallel paths for the signal. In the case of a distributed amplifier, each pathway provides some gain; therefore, the whole amplifier is capable of providing a higher gain-bandwidth product than a conventional amplifier.

Other Distributed Circuits

While the distributed amplifier discussed in the previous section is a good illustrative example of a distributed system, it is not the only practical one. The delay introduced by the distributed amplifier can be used to create an oscillator by feeding its output back to its input [4, 6, 7]. The oscillation frequency is determined by the round-trip time delay, that is, the time it takes the wave to travel through the transmission lines and get amplified by the transistors, as shown analytically in [4].

The frequency of the distributed oscillator should be tunable to be useful in a phase-locked system. As mentioned earlier, the oscillation frequency is inversely proportional to the total delay and hence the total length of transmission lines. This property leads to a frequency tuning approach based on changing the effective length of the transmission lines.

Although the physical length cannot be changed, the effective length can be varied, that is, we can control and change the path traveled by the waves. The basic concept illustrated in Fig. 4 shows a replacement for each single-transistor segment of the distributed amplifier of Fig. 3b. The gates of transistors M_a and M_b are connected to the same point on the gate line while separated on the drain line by distance, l. The frequency of oscillation can be changed by altering the gain of each transistor through controlling the ratio of the two currents, I_1 and I_2. To understand this tuning method, consider two extreme cases. First, when M_a has the maximum gain and M_b is off, the signal travels along the solid path. In the second extreme case where M_b has its maximum gain and M_a is off, the signal travels along the dashed path. The latter is shorter than the former by l. Therefore, the round-trip time delay and hence the oscillation frequency can be varied by adjusting the ratio of the bias currents. This concept can be seen using the racetrack analogy of Fig. 5. Here the signals traveling on the input and output lines are analogous to two runners on two tracks running side by side to be able to pass a torch at all times. They run at a constant speed, since the propagation velocity of the waves on both lines are fixed. The time it takes them to complete a lap (oscillation period) can be changed by introducing symmetrical shortcuts for both of them and controlling what percentage of the time they go through the shortcuts. This technique has been used to demonstrated a 10 GHz distributed voltage controlled oscillator (DVCO) using a 0.35 μm CMOS transistor with a tuning range of 12 percent and phase noise of –114 dBc/Hz at 1 MHz offset [4]. In addition to higher frequency of oscillation, the DVCO provides lower sensitivity to process variations and better frequency stability.

Yet another example of distributed systems is that of the *distributed active transformer* (DAT) power amplifier [8]. The design of a fully integrated power amplifier with reasonable output power, efficiency, and gain has been one of the remaining major challenges in today's pursuit of a single-chip integrated transceiver. Although several advances have been made in this direction, a watt-level truly fully integrated CMOS power amplifier had not been demonstrated using the traditional power amplifier design techniques[4] before the recent introduction of the distributed active transformer DAT [8].

Two main obstacles in the design of a fully integrated power amplifier are the low breakdown voltages of transistors and the high loss of passive components. The low breakdown voltage limits the voltage swing at the output node which in turn lowers the produced output power. The high passive loss reduces amplifier's power efficiency by dissipating the generated power in the signal path. These problems are exacerbated in most commonly used CMOS process technologies, as the MOS transistor's minimum feature size is continuously scaled down for faster operation, resulting in lower substrate resistivity and smaller breakdown voltages.

The DAT power amplifier [8] uses a distributed method to perform impedance transformation and power combining simultaneously to achieve large output power while maintaining acceptable power efficiency. It overcomes the low breakdown voltage of short-channel MOS transistors and alleviates the substrate loss problems by providing the power gain through multi-

4 To date, the highest power levels achieved with fully integrated amplifiers in standard silicon are on the order of 100 mW [9].

ple similar stages and signal paths. Figure 6 shows the essential features of the DAT.

The DAT consists of multiple distributed push-pull circuits in a polygon geometry, as seen in the square geometry of Fig. 6. Each side of the square is a single push-pull amplifier consisting of a load transmission line, two transistors, and input matching transmission lines. The push-pull structure creates a virtual ac ground at the supply node that makes it unnecessary to use choke inductors and large on-chip bypass capacitors at supply. This particular positioning of the push-pull amplifiers makes it possible to use a straight wide metal line as the drain inductor. These slab inductors provide natural low-resistance paths for the dc current to flow from the supply to the drain of the transistors, as shown in Fig. 6.

A power splitting network in the center of the DAT (not shown in Fig. 6) provides the differential input to the transistor pairs in the corner. By driving the two adjacent transistors of two different push-pull amplifiers in opposite phases, we can create a virtual ac ground in each corner of the square. This is an essential feature for realizing a lumped inductor characteristic using a transmission line whose two ends are at two different physical locations.

The four ac-coupled load transmission lines are used as the primary circuit of a magnetically coupled transformer to combine in series the output power of these four push-pull amplifiers and match their small drain impedance to a 50 Ω load. These four push-pull amplifiers, driven by alternating phases, generate a uniform circular current at the fundamental frequency around the square, resulting in a strong magnetic flux through it. A one-turn metal loop inside the square is used to harness this alternating magnetic flux and act as the transformer secondary loop. This is where multiple signal paths converge. Using the distributed active transformer, a fully integrated watt-level power amplifier was demonstrated in a standard CMOS process technology for the first time [8]. It provides 2 W output power on a 2 V power supply using a 0.35 µm CMOS transistors, while achieving a power added efficiency (PAE) of 42 percent. The distributed nature of the DAT structure reduces the sensitivity of the power amplifier's efficiency to the substrate power losses while providing a large overall output power using low-breakdown-voltage MOS transistors. The strong electromagnetic coupling between multiple signal paths in a DAT necessitates an analysis and design approach spanning architecture, circuits, device physics, and electromagnetics.

These three examples — distributed amplifier, DVCO, and DAT power amplifiers — demonstrate some of the basic concepts of distributed integrated circuit design. The combination of multiple distributed signal paths working in harmony and a design approach covering several levels of abstraction allow us to achieve higher frequencies of operation, higher power and efficiency, while creating a more robust system.

CONCLUSIONS

The above discussion and examples demonstrate that distributed circuits can serve as effective means to implement high-performance integrated circuits. This methodology is particularly useful when the desired frequency of operation is close to the cut-off frequency of the transistors. The use of multiple parallel signal paths working in synchronization to distribute parasitics and divide the functionality evenly results in significant improvements in the design. Ironically, this systematic spreading of the gain and parasitics among the components is yet another example of the prevailing divide-and-conquer strategy.

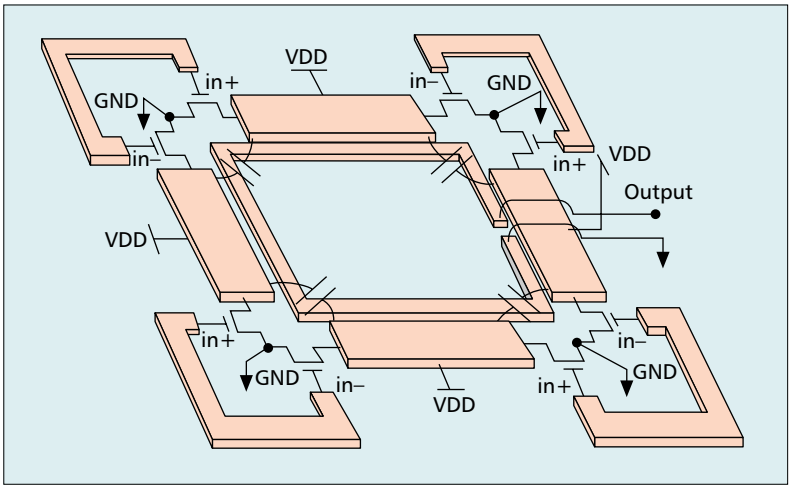

Figure 6. *The basic structure of the DAT power amplifier depicting only the top metal layer for simplicity. The input power splitting network is not shown.*

ACKNOWLEDGMENTS

The author would like to thank all members of the Caltech High-speed Integrated Circuits (CHIC) group: D. Ham, H. Wu, H. Hashemi, R. Aparicio, I. Aoki, C.J. White, B. Analui, X. Guan, A. Komijani, S. Kee, E. Afshari, and A. Sridhar, as well as Prof. D. Rutledge of Caltech for helpful discussions and their contributions to this article.

REFERENCES

[1] W. Jutzi, "A MESFET Distributed Amplifier with 2 GHz Bandwidth," *Proc. IEEE*, June 1969, pp. 1195–96.
[2] W. S. Percival, British Patent Specification, No. 460,562, applied for July 24, 1936.
[3] E. L. Ginzton, W. R. Hewlett, J. H. Jasberg, and J. D. Noe, "Distributed Amplification," *Proc. IRE*, vol. 36, Aug. 1948, pp. 956–69.
[4] H. Wu and A. Hajimiri, "Silicon-Based Distributed Voltage Controlled Oscillator," *IEEE J. Solid-State Circuits*, Mar. 2001.
[5] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge, U.K., Cambridge Univ. Press, 1998.
[6] L. Divina, and Z. Skvor, "The Distributed Oscillator at 4 GHz," *IEEE Trans. Micro. Theory Tech.*, vol. 46, no. 12, Dec. 1998, pp. 2240–43.
[7] B. Kleveland *et al.*, "Monolithic CMOS Distributed Amplifier and Oscillator," *IEEE ISSCC Dig. Tech. Papers*, Feb. 1999, pp. 70–71.
[8] I. Aoki *et al.*, "A 2.4-GHz, 2.2-W, 2-V, Fully-Integrated CMOS Circular-Geometry Active-Transformer Power Amplifier," *IEEE CICC Dig. Tech. Papers*, San Diego, CA, May 2001.
[9] Y. J. E. Chen *et al.*, "RF power Amplifier Integration in CMOS Technology," *IEEE MTT-S Digest*, vol. 1, Boston, MA, June 2000, pp. 545–48.

The strong electromagnetic coupling between multiple signal paths in a DAT necessitates an analysis and design approach spanning architecture, circuits, device physics, and electromagnetics.

Additional Reading

[1] A. Hajimiri and H. Wu, "Analysis and Design of Silicon Bipolar Distributed Oscillators," *IEEE VLSI Symp. Dig. Tech. Papers*, June 2000, pp. 102–5.

Biography

Ali Hajimiri (hajimiri@caltech.edu) is an assistant professor of electrical engineering at California Institute of Technology (Caltech). He received his B.S. degree in electronics engineering from the Sharif University of Technology, and his M.S. and Ph.D. degrees in electrical engineering from the Stanford University, Stanford, CA, in 1996 and 1998, respectively. He was a design engineer with Philips Semiconductors, where he worked on a BiCMOS chipset for GSM and cellular units from 1993 to 1994. In 1995 he was with Sun Microsystems, where he worked on the UltraSPARC microprocessor's cache RAM design methodology. In 1997 he was with Lucent Technologies (Bell Labs), Holmdel, New Jersey, where he investigated low-phase-noise integrated oscillators. He joined the faculty of Caltech in 1998, where his research interests are high-speed and RF integrated circuits. He is a coauthor of *The Design of Low Noise Oscillators* (Kluwer, 1999) and has received several U.S. and European patents. He is an associate editor of *IEEE Transactions on Circuits and Systems* (TCAS): Part-II and a member of the Technical Program Committees of the International Conference on Computer Aided Design (ICCAD). He has also served as guest editor of *IEEE Transactions on Microwave Theory and Techniques* and on the Guest Editorial Board of *Transactions of Institute of Electronics, Information and Communication Engineers* of Japan (IEICE). He was the Gold medal winner of the National Physics Competition and the Bronze Medal winner of the 21st International Physics Olympiad, Groningen, the Netherlands. He was a co-recipient of the International Solid-State Circuits Conference (ISSCC) 1998 Jack Kilby Outstanding Paper Award and winner of the IBM faculty partnership award.