Application of an SOI 0.12-μm CMOS technology to SoCs with low-power and high-frequency circuits

Systems-on-chips (SoCs) that combine digital and high-speed communication circuits present new opportunities for power-saving designs. This results from both the large number of system specifications that can be traded off to minimize overall power and the inherent low capacitance of densely integrated devices. As shown in this paper, aggressively scaled silicon-on-insulator (SOI) CMOS is a promising technology for SoCs for several reasons: Transistor scaling leads to active power reduction in the sub-50-nm-channel-length regime, standard interconnect supports the high-quality passive devices essential to communications circuitry, and high-speed analog circuits on SoI are state of the art in terms of both performance and power dissipation. We discuss the migration of a complete digital circuit library from bulk to SOI to prove that SOI CMOS supports ASIC-style as well as fully custom circuit design.

1. Introduction

Many of the power-management schemes currently used in integrated circuits involve adding a power-control system to what are essentially standard circuit building blocks. These power-control networks supply power to the blocks that limit throughput at any given moment and remove it from the blocks that are not supplying useful output. Examples of such power-control schemes include clock gating and variable power supplies. However, when an integrated circuit is a true system-on-a-chip (SoC), encompassing not only digital and memory blocks but also high-frequency analog blocks for wired and wireless communication, power savings become more than a question of the computational efficiency of digital logic. In communications circuits, the power budget depends on such elements as noise levels, isolation, transmission efficiency between subcircuits, and the losses in passive components. SoCs allow the reduction of total system power because the performance and communication specifications of a system can be met with subcircuits and transmission networks that are inherently more energy-efficient than discrete units in a package.

Systems are becoming increasingly integrated, and it is becoming obvious that SoCs are indeed feasible. The first example of the combination of analog and digital blocks was the integration of clock generator phase-locked loops (PLLs) with processors and memory. CMOS circuits with even more analog and radio frequency (RF) function are now being introduced. For example, products such as 5-GHz-wireless-LAN, 1.9-GHz-GSM-cellular, and 10-Gb/s-SONET transceivers [1–3] are being fabricated in 0.18-μm and 0.13-μm technologies, though without integration of the baseband digital signal processor or the microprocessor. As shown in Figure 1(a), the receiving

Copyright 2003 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the Journal reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to reproduce any other portion of this paper must be obtained from the Editor.

IBM J. RES. & DEV. VOL. 47 NO. 5/6 SEPTEMBER/NOVEMBER 2003

J.-O. Plouchart et al.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
When combined with a microprocessor, a transceiver makes possible the highly connected, highly intelligent world of embedded electronics envisioned by many. In the integration race, the next step might be the integration of these two functions on the same chip (Figure 1). In the future, wireless and wired transceivers will be seen simply as I/O for microprocessors. As integration continues, we can begin to envision a one-chip solution. This will have a profound and beneficial effect on system cost, power, and size. There are, of course, many barriers to the one-chip solution: Sensitive analog circuits can be disturbed by noise-producing digital logic, the conflicting requirements of analog/RF circuits [4] and digital circuits demand the integration of multiple device types, and there will always be competition from system-in-a-package solutions.

In this paper we discuss silicon-on-insulator (SOI) technology, the platform of choice to integrate digital and high-speed analog functions owing to the 10X-lower parasitic capacitance to the substrate of active SOI devices and its inherent higher efficiency and isolation when compared with bulk silicon technologies. SOI ultralarge-scale integration (ULSI) capabilities and the integration of DRAM and SiGe bipolar transistors on SOI have previously been reported [5–7]. Some aspects of low-power digital design on SOI have also been reported [8, 9]. We show here that, as the number of circuit functions supported by a technology increases and as the performance of those elements increases, the opportunity to save power grows. We first discuss how scaled SOI CMOS provides increased performance and reduced active power in the sub-50-nm-channel-length regime. We then present passive devices integrated with standard interconnect and discuss how they support low-voltage, low-power analog blocks. Next, we discuss circuits designed in this technology and present measurements of high-speed analog circuits and digital signal processing (DSP) building blocks that are state of the art in power dissipation and performance and a circuit library that encompasses low-active-power digital and mixed-signal blocks and allows simple bulk-to-SOI migration. This paper presents new and recently published research and development work done across several IBM organizations on SOI SoC.

2. Power saving enabled by high-performance SOI FETs

FET scaling and active power

From generation to generation, transistor gate-length and oxide-thickness scaling have successfully reduced the active power dissipation of gates switched at a given fixed
frequency. This reduction has depended on transistor characteristics (for example, input capacitance, drive current, and operating voltage) following the simple scaling trends first discussed by Robert Dennard [11]. In the future, continued reduction of active power dissipation will be difficult if parasitic elements, such as fringe capacitances and series resistances, diminish the benefits of scaling.

One way to determine whether scaled transistors obey traditional scaling trends is to measure their small-signal parameters at high frequency. These parameters are measured at actual operating frequency (1 to 10 GHz) and give more information about transistors than gate delay alone, including important performance- and power-determining parameters such as input capacitance and transconductance. It has been shown that the gate delay of scaled CMOS technology decreases continuously as transistor gate length decreases in the 45-nm to 70-nm region [12]. However, such behavior would be observed even if punchthrough and poor gate control occurred at the shortest gate lengths. For active power to decrease with transistor scaling, the transconductance and other small-signal parameters must improve as gate length decreases.

Most publications on the high-frequency properties of CMOS transistors provide data on overall figures of merit, such as the unity current-gain frequency ($f_T$) and unity power-gain frequency ($f_{max}$), but they do not present data on more particular transistor parameters such as transconductance ($g_m$) and input capacitance ($C_{in}$) [13, 14]. These publications show that $f_T$ and $f_{max}$ are improving with gate-length scaling when one's view extends across technologies and technology generations. However, these publications do not show evidence of well-behaved small-signal parameters across a range of short (<100 nm) channel lengths. Indeed, one publication that does provide $g_m$ and $C_{in}$ data shows a decrease in transconductance at gate lengths below 80 nm [15]. In the following paragraphs, we present small-signal parameter data which suggest that active power will decrease as FETs are scaled [10].

To extract the SOI FET small-signal parameters, we first measured the scattering parameters (S-parameters) using a vector network analyzer. An S-parameter is defined to be the ratio of the scattered voltage wave to a wave incident on an FET. By manipulating the S-parameters, the current gain ($H_{21}$) can be computed, as well as the impedance ($Z_{in}$) or admittance ($Y_{in}$) parameters. We measured the S-parameters of four n-FETs from a 0.12-μm partially depleted SOI CMOS technology [16] ranging in gate length from 47 nm to 72 nm. Each device consisted of 64 fingers, each 1.24 μm long. The parameters $f_T$, $g_m$, and $C_{in}$ were extracted from their respective de-embedded small-signal parameters—$Mag(H_{21})$, $Real(Y_{in})$, and $Imag(Y_{in})$—in the frequency range of 2 GHz to 20 GHz, over which each parameter had the ideal frequency dependence (proportional to $1/f$, constant, and proportional to $f$, respectively). The measurements were made with a constant dc bias at a relatively high power density. Therefore, the data (particularly the $g_m$ data) is influenced by self-heating. The results are shown in Figure 2. None of the small-signal parameters reach a limiting value at the smallest gate length of 47 nm. The input capacitance shows a tendency to saturate at even lower gate lengths as a result of finite gate-fringing capacitance. The transconductance shows no sign of saturation, indicating that the gate has good control of the channel and that the effective velocity of electrons in the channel increases with increasing lateral electric field, even at high drain voltage (1.2 V). The $f_T$ of 196 GHz at the shortest gate length is the highest ever reported for a silicon FET with acceptable short-channel effect [15].

No discussion of the small-signal parameters of FETs would be complete without a discussion of the maximum frequency of oscillation, $f_{max}$. This is one of the few small-signal parameters that reflect the input resistance of the transistor gate, which can decrease switching speed at a given level of power dissipation. Perhaps the most difficult parameter to measure accurately is $f_{max}$, because it requires accurate measurement of a very low input resistance (~1 Ω) that is in series with the high-impedance input capacitance (~90 fF). The theoretical resistor/capacitor (RC) time of the input corresponds to a frequency above 1 THz, which means that the input resistance is not resolvable in the frequency range accessible to basic equipment (<20 GHz) and, in higher-frequency systems, that it can be distorted by small calibration errors and contact resistance.

We extracted the unilateral power gain from our S-parameter data and found that it had greater than the ideal 20-dB-per-decade slope over all frequencies in the 2-GHz to 110-GHz range. No simple FET model can produce this behavior. Thus, we refrain from stating a measured $f_{max}$ value. However, at 20 GHz, a frequency at which calibration is usually reliable and the unilateral gain can be measured precisely, every FET has a unilateral power gain greater than 20 dB, indicating that $f_{max}$ is probably greater than 200 GHz for every FET. Calibration techniques and the dependence of $f_{max}$ on transistor layout must be studied further to make accurate $f_{max}$ measurement possible.

One parameter that may be vulnerable to scaling is the resistance of the silicide on a transistor gate. In [14] it was reported that gate sheet resistance is an increasing function of gate length at drawn gate lengths near 180 nm. It was also shown through process simulation that silicide resistance is the dominant input loss mechanism in the same gate-length regime. We measured the gate resistance...
on FETs with high, easily measurable total gate resistance (−150 Ω) to investigate the gate-length dependence of silicide sheet resistance at gate lengths below 100 nm. The results are shown in Figure 2(d). The measured resistance is effectively constant below 20 GHz, as expected, reflecting accurate measurement. The rolloff in the gate resistance above 20 GHz is due to either parasitic inductance or inaccurate calibration (the gate capacitance of those structures was de-embedded). The sheet resistance is 6.3 Ω/μm at a gate length of 77 nm and rises to 7.3 Ω/μm at a gate length of 55 nm. This sheet resistance increase is moderate enough not to be a performance limiter and to allow aggressively scaled FETs to have very high f_T.

We measured the active power dissipated by an actual CMOS circuit (a ring of unloaded inverters) for further confirmation that active power is reduced as transistor gate length is scaled. Figure 3 shows the dependence of normalized active energy per switch on gate delay for unloaded inverters of three different gate channel lengths. The normalized energy is calculated as

\[ \tau (I_{\text{dd,active}} - I_{\text{dd,quiescent}}) W_s / (W_n + W_p), \]

where \( \tau \) is the gate delay and \( W_n \) and \( W_p \) are the n-FET and p-FET widths, respectively. The delay and energy were modulated by tuning the power-supply voltage, \( V_{dd} \), between 1.2 V and 0.6 V. The active energy decreases continuously with decreasing gate length at every gate delay in the range measured. This range corresponds to a variation in performance of a factor of 4, which is greater than the typical performance tuning range of power-saving circuitry that selectively diminishes the power-supply voltage of noncritical paths (voltage scaling) [17]. Therefore, gate-length scaling supports typical powersaving circuit architectures because it leads to active power reduction in both high-performance and low-performance regimes. The active energy of 0.44 fJ/μm2 dissipated by the ring at the shortest channel length is the lowest ever reported at a gate delay of 25 ps.
FET RF and low-frequency noise

Noise has a direct impact on the power dissipation of a communication system because the noise level determines the minimum signal power needed to maintain the signal-to-noise ratio. As long as ample gain (~10 dB) is maintained, lowering the noise of active devices allows the reduction of signal power, thereby reducing overall power consumption. The RF noise figure and associated gain are the most important figures of merit for the detection and processing of low-amplitude signals. We measured an n-FET with an $L_{pol} = 65$ nm to have a minimum noise figure, $NF_{min}$, less than 2 dB up to 26 GHz (see Figure 4). We also measured an associated gain of 19 dB at 2 GHz. This data is the best ever reported for a silicon technology and is similar to that of a state-of-the-art pseudomorphic high electron mobility transistor (PHEMT). For example, at 12 GHz, the PHEMT ATF-36077 [18] exhibits a 0.5-dB $NF_{min}$ and an associated gain of 12 dB, compared with 1 dB and 14.5 dB, respectively, for our n-FET.

Low-frequency noise is an important figure of merit for analog circuit designers because it affects the jitter of voltage-controlled oscillators (VCOs) and the charge pump of PLLs. The high nitrogen content of the thin gate oxide of scaled CMOS is known to increase low-frequency noise [19]. Designers can take advantage of the thick ($T_{ox} = 2.9$ nm) dual oxide offered in our technology, which has lower 1/f noise than the standard oxide owing to its different nitrogen distribution. As shown in Figure 5, the measured 1/f noise of a thick-oxide n-FET with its body tied to ground is 3× lower than the noise reported for the thick-oxide (7.0-nm) n-FET of an 0.18-μm bulk technology [20]. The comparison is fair because the 4× difference in active area is compensated by the 2× difference in current (see Figure 5 for device dimensions and currents). This demonstrates that oxides with low 1/f noise and aggressively scaled oxides can be integrated despite the high nitrogen content of the scaled oxides. In coming generations, a FET with acceptable 1/f noise performance should be continuously available to designers of noise-sensitive circuits.
The drain-body diode current of floating-body SOI FETs is known to be a distinct source of low-frequency noise [21]. The shot noise of the diode current induces noise in the body voltage which, in turn, varies the channel current. This noise mechanism has a Lorentzian frequency dependence that is markedly different from a 1/f dependence, making it distinguishable in measured noise spectra. Diode-induced noise can be suppressed by using devices with body contacts. The noise shown in Figure 5 has no Lorentzian signature, showing successful suppression of diode-induced noise. This shows that our SOI technology offers FETs free of SOI-specific noise mechanisms, with 1/f noise levels as good as or better than those of bulk FETs.

3. Integration of high-performance passive devices in a standard BEOL

Scope

For the design of wired or wireless transceivers, high-performance passive devices are used to build matching circuits, filters, couplers, transformers, or resonators. The integration of passive devices is critical for SoCs based on CMOS technology. Unfortunately, there is no alternative method of implementing passive devices with performance comparable to that of discrete passive devices [22], and sizable real estate is required for the integration of the discrete passive devices onboard. However, the feature size of CMOS technology is becoming smaller, and the back-end-of-line (BEOL) and front-end-of-line (FEOL) options are now very diversified. There are many attributes requested for passive devices, such as linearity, temperature stability, tolerance, matching, and so on, but one of the most important parameters for reactive elements is losses. High-performance passive devices have low losses or a high quality factor. The quality factor, usually termed \( Q \), is defined at a frequency \( f \) by

\[
Q = 2\pi f \quad \text{(energy stored/average power dissipated)},
\]

where \( Q \) is dimensionless and is proportional to the ratio of energy stored to energy lost per unit time. Intuitively, we can understand that this is a very important factor for low-power applications, and that a low-power technology will try to maximize \( Q \) by reducing the energy lost.

Resonant circuits, such as a capacitor (\( C \)) in parallel with an inductor (\( L \)), are used in many circuits (e.g., amplifiers, mixers, or oscillators). It can be shown that an \( L, C \) parallel circuit is equivalent at the resonance to a resistor,

\[
R = Q \frac{L}{C}.
\]

For a certain voltage supply (\( V \)), the current consumption is

\[
I = \frac{V}{Q \sqrt{L/C}}.
\]

Therefore, by maximizing \( Q \), we can decrease the current consumption. Passive device losses clearly have a direct and negative impact on efficiency and chip power consumption. In this section, we describe the source of losses and high-performance inductors and capacitors.

Source of losses

Sources of losses are ohmic and radiation losses in metals, dielectric losses, and eddy currents in semiconductors. Unfortunately, the silicon substrate usually used in silicon technologies is very lossy. The substrate quality coefficient can be derived from the dielectric relaxation equation

\[
F_{\text{rel}} = \frac{1}{2\pi\rho\varepsilon_0\varepsilon},
\]

where \( \rho \) is the substrate resistivity in \( \Omega\cdot\text{cm} \), \( \varepsilon_0 \) is the dielectric constant of the vacuum \( (8.854 \times 10^{-16} \text{ F/cm}) \), and \( \varepsilon_r \) is the dielectric constant of the substrate \( (\sim 12 \text{ for silicon substrate}) \). The dielectric relaxation frequency is also equal to the cutoff frequency of the substrate, \( 1/(2\pi \rho\varepsilon_0\varepsilon) \), where the substrate is modeled by a resistor \( R_s \) in parallel with a capacitor \( C_s \). Because the \( Q \) of the substrate at a frequency \( F \) is \( 2\pi R_sC_sF \), one can derive that it is also equal to \( F/F_{\text{rel}} \). For mixed-signal technologies,
substrate resistivities as high as 15 Ω-cm are used; this is equivalent to a dielectric relaxation frequency of 10 GHz. The Q at 1 GHz of the substrate is therefore only 0.1. This is a very important limitation factor when one wants to integrate high-Q inductors on chip. An obvious approach to raising the Q of the substrate is to increase its resistivity. This is not always feasible in a bulk technology, because bulk junction isolation requires a conductive substrate to shunt carriers that might otherwise transiently forward bias the isolating junctions (latchup). An SOI technology has dielectric isolation that is independent of the substrate; therefore, the substrate resistivity is a free parameter. A high-resistivity silicon (HRS) substrate can be used to add value to the passive devices by boosting their performance.

**High-performance inductors**

There are two main inductor specifications, Q and inductance density, which are inversely proportional to each other. Therefore, we need to optimize inductor designs for different applications. We classify these into two types of inductors, which are discussed with their respective optimized design schemes. To boost inductor Q, we use high-resistivity substrate (HRS > 100 Ω-cm) instead of regular-resistivity substrate (RRS = 12 Ω-cm). Two different categories of on-chip inductors target different applications: single-turn parallel multilayers (STPML) and multi-turn series multilayers (MTSML).

**STPML inductors**

Figure 6 shows a cross section of an eight-level copper-metal interconnection of 0.12-μm SOI CMOS technology [23]. The substrate resistivity can be changed from 12 Ω-cm (RRS) to 300 Ω-cm (HRS). The BEOL consists of groups of levels of common pitch and thickness, where the pitch and thickness of each group are an integer multiple of the minimum (hence, the group names 1X, 2X, and 4X). Figures 7(a) and 7(b) show a metal-stack technique to improve conductivity for STPML inductors [24]. We can stack up to four metals, M8 to MS, and via contacts are used between metal layers. The maximum copper thickness with four stacked metals is 3.6 μm for the metal and 3.8 μm for the vias. The dc resistance is dramatically reduced via parallel metal stacking.

Figure 7(c) shows the measured Q [24]. The data prove that the parallel metal-stack geometry improves the Q factor, but the self-resonant frequency is slightly reduced because of the substrate effect. The highest Q is about 52.
at 4 GHz via the use of a parallel metal stacked inductor on an HRS substrate. The inductor density of 6 fH/μm² is the lowest. Figure 7(c) shows that, owing to HRS, inductor Q values can be improved up to 80%.

STSML and MTSML inductors
Figures 8(a) and 8(b) show the series-connection geometry of the STSML inductor [24], which emphasizes inductance density and is a tradeoff between Q and density. The end of M8 is connected to the end of M7, and the front of M7 goes to the front of M6. The end of M6 is connected to the end of M5. After all series metal connection, the two terminals of this inductor are the front of M8 and the front of M5. The total inductor length can be extended four times longer than that of the parallel metal connection.

Figure 8(c) shows the measured Q for STSML and MTSML inductors [24]. For high-inductance density, we can use as many metal layers and turns as possible. The MTSML3T8 inductor produces the highest inductance density, but Q will be relatively damped, as shown in Figure 8(c). However, the STSML1T4 can improve Q and self-resonant frequency. Measured inductance values are 42 nH for MTSML3T8 and 2 nH for STSML1T4. The inductance density of MTSML3T8 is a record 320 fH/μm². This inductance density is at least 50X higher than that of the STPML inductor. MTSML inductors can be used as choke coils for RF amplifiers and in broadband applications.

Figure 9 shows a comparison of reported state-of-the-art inductors [25–32]. It demonstrates that a standard microprocessor SOI BEOL can compete very well with micromachined inductors or inductors requiring special GaAs or sapphire substrates as well as thick dielectric and metal. Note that these specialized BEOLs require extra processing steps and are therefore much more costly.

High-performance capacitors
The requirements for high-performance capacitors are high capacitance density, low capacitance tolerance, and high Q. These three factors can be traded off and are dependent on capacitor geometries. New CMOS technologies are currently offering many options for BEOL as well as FEOL to improve circuit integration and performance. Here we focus on two types of high-performance capacitors: the metal–insulator–metal (MIM) capacitor, suitable for low-tolerance requirements, and the vertical-parallel-plate (VPP) capacitor, suitable for high-Q applications.

MIM capacitor
Figure 10(a) shows a typical layout for a MIM capacitor [23]. The MIM capacitor is built between M5 and M6 with
special mask layers for the bottom and for the top. The silicon nitride dielectric is deposited between the layers to increase capacitance. The core capacitance density achieved is 1 fF/μm². Depending on the material used for the top and the bottom plates, the Q of the MIM capacitor will be damped. However, instead of using a single device structure, as in Figure 10(a), we can improve the capacitor Q by designing a multi-finger structure. Figure 10(b) shows the difference in Q between a single-finger device and a multi-finger device of the same total capacitance. The quality factor of a small capacitor finger, QB, is higher than that of a large single-finger device, QA, because a large device has higher access resistance than a small finger. If many small fingers are connected in parallel to achieve the same total capacitance, the Q of the multi-finger capacitor, QM, is ideally the same as QB. Therefore, the multi-finger geometry can boost the Q of equivalent capacitors.

Figure 10(c) shows a multi-finger MIM capacitor in which eight single-finger capacitors are connected in parallel [23]. As we can see in the microphotograph, we must trade off capacitance density and Q. The multi-finger structure can improve capacitor Q, while the capacitance density is degraded as a result of the required extra interconnect area. We must then recalculate the capacitance density by taking into account the extra wiring area needed, such as effective area I and effective area II, as shown in Figure 11(a). The effective MIM capacitance

![Figure 9](https://example.com/figure9.png)

**Figure 9**
Inductance as a function of peak Q [24]. (The numbers near the symbols are the reference number and the peak-Q frequency in GHz.)
density can easily be reduced by a factor of 2 to 3. The MIM capacitor tolerance is ±15%. We can apply MIM capacitors for RF matching circuits as accurate reactive capacitors.

**VPP capacitor**

A MIM capacitor is a lateral device, and it is not easy to implement symmetrical structure. In general, asymmetrical passive devices are not recommended because their use usually produces layout limitations. Therefore, we proposed a VPP capacitor, taking advantage of the many combinations possible in advanced BEOL. A three-dimensional geometry view of a VPP capacitor is shown in Figure 11(a) and a microphotograph in Figure 11(b) [23]. There are three different metal connections noted: Q1, Q2, and Q3. Q1 is composed of 1X metals (M1, M2, M3, and M4), Q2 is 2X (M5 and M6) metals, and Q3 has 4X metals (M7 and M8). Capacitance C3 is formed of 4X metals (M7 and M8) and the top via contact (V7). C3 has the lowest capacitance density because of the wider distance between 4X metals. C1 is generated with 1X metals and the bottom via contacts (V1, V2, and V3). C1 has the highest capacitance density, since these metals can be placed more closely. C2 is formed by the 2X metals and via contacts (V4).

VPP capacitors use all combinations of metals to boost the capacitance density. Vias are used for the contacts between metals, and theses contacts significantly increase the capacitance density. The VPP capacitance density increases with the number of interconnections and is now comparable with MIM capacitor density. As shown in Figure 11(c), we measured the \( Q \) of VPP and MIM capacitors. The 1.1-pF VPP capacitor occupies a 625-\( \mu m^2 \) area. Therefore, the capacitance density is 1.76 \( \text{fF/}\mu\text{m}^2 \). On the other hand, the 0.8-pF MIM capacitor occupies an 800-\( \mu m^2 \) area, and its capacitance density is 1.0 \( \text{fF/}\mu\text{m}^2 \). The VPP achieves at least 80% higher capacitance density than the MIM capacitor, which requires the use of multi-finger devices to improve \( Q \), but the capacitance density is decreased because of the extra wiring required. Figure 11(c) also shows that MIM capacitor \( Q \) is lower than for the VPP capacitor beyond 2 GHz. The VPP capacitor can be fabricated without any extra integration process, while the MIM capacitor requires extra masks and process steps. The \( Q \) of the VPP can be increased by layout optimization, and another of its advantages is capacitor symmetry.
Summary of passive devices
For SoC applications, a large variety of high-performance passive devices are required. We have described high-performance inductors and capacitors in this section and showed that many different types of passive devices can be integrated in a modern SOI microprocessor BEOL. We have also proved that passive devices are improved with the use of high-density BEOL and HRS. Owing to the use of HRS, vertical structure can be built with less penalty from substrate losses. Also, inductor Q and inductor density can be traded off with serial or parallel metal stacking.

4. Circuit examples and circuit isolation

Low-power, low-jitter clock generation
Key to SoC clock generation are low-power, low-phase-noise VCOs that use on-chip inductors and varactors. A tuning range of more than 15% is usually required to cover the frequency band and to compensate for temperature and process variations. Beyond a 20-GHz oscillation frequency, wide-tuning-range LC CMOS VCOs are difficult to design. For example, a 51-GHz VCO in 0.12-μm CMOS with less than 3% tuning range was reported [33]. A higher tuning range of 9% on a 1.5-V supply has been reported at 40 GHz on SOI [34]. We report here a 15% tuning range for a 40-GHz VCO on HRS at 1.5 V, demonstrating for the first time the potential manufacturability of such a high-frequency VCO. A VCO that oscillates from 35.2 GHz to 40.9 GHz from 0 V to 1.5 V across the varactor is shown in Figure 12. In contrast to the work of [33], where only n-MOS transistors are used to generate negative resistance, a more symmetrical cross-coupled inverter using n-MOS and p-MOS is used, as illustrated in Figure 12(a). This circuit demonstrates the high-speed and low-parasitic capabilities of SOI technology. The power consumption is only 12 mW in the core and 5.7 mW in the buffer from a 1.5-V supply. The measured phase noise is ~90 dBc of the 35.6-GHz carrier, as shown in Figure 12(c). The measured output power is more than ~20 dBm across the full tuning range, as shown in Figure 12(b).

High-speed, low-power CML
For high-speed digital processing, current mode logic (CML) is the preferred architecture. The static frequency divider-by-two function is usually the slowest because of the feedback loop used. Figure 13(a) shows the block diagram and latch of the implemented static CML frequency divider-by-two [35]. Resistor loads of 400 Ω were used for the latch, and no inductive peaking was used to extend the bandwidth frequency. Figure 13(b) shows the divider sensitivity as a function of frequency for different voltage supplies [35]. Because three transistors are stacked between $V_{dd}$ and ground, we can use voltage supplies as high as 2 V without compromising the reliability of the transistor. A maximum operating frequency of 30 GHz is achieved for a power consumption per latch of 6 mW from a 2-V supply. At a 1.5-V supply, a maximum operating frequency of 28.6 GHz is achieved, for a power consumption per latch of 3.83 mW. A CMOS
0.12-μm-frequency divider-by-two using the same CML latch architecture without inductive peaking operates at up to 18.5 GHz [36], for an input power of more than 10 dBm, with a power consumption of 13.5 mW per latch from a 1.5-V supply. At a 1-V supply, the SOI frequency divider is still operational up to a maximum frequency of 25 GHz, with a power consumption of 1.35 mW per latch. The SOI CML frequency divider is also more sensitive at maximum frequency than the bulk version [36].

One way to compare the power and speed performance of different circuit dividers is to compute the power–delay product. At 1 V, the SOI divider exhibits a record energy per gate of 13.5 fJ, assuming two gate delays per flip-flop and an equivalent complexity of four logic gates [37]. Figure 14 shows a comparison of state-of-the-art static dividers in several technologies [35, 38–41]. To the authors’ knowledge, the lowest static divider energy reported for any technology at an operating frequency higher than 20 GHz is that of 0.12-μm SOI technology. The closest energy is 24 fJ for an AlInAs HBT technology, which is 78% higher than that reported in this work. The mechanism for energy reduction is very different for an SOI CMOS technology than for an HBT compound technology. Owing to the threshold- and supply-voltage scaling, a low operating voltage supply can be used. This allows a dramatic reduction of switching energy. If we compare with bulk CMOS, the lower parasitic capacitance offered by SOI technology is an important factor in reducing power consumption. This demonstrates the speed and power advantages of SOI technology for the CML latches used extensively for RF and high-speed communications. These results, combined with the ultralarge-scale integration (ULSI) capabilities of the technology, are very promising for the integration of multiple high-speed serial links on the same chip. This integration could lead to the aggregation and processing of greater amounts of data than ever before.

**High-speed, low-power digital library**

We have described above how the active and passive elements of an SOI technology support the various functions of an SoC. However, SoCs on SOI will not be truly producible until designers can access digital libraries that enable the design of millions of gates of power-efficient digital logic. The design and production of such a library, the first on SOI in the industry, is currently underway at IBM for 0.12-μm technology. The library supports a wide variety of digital logic gates, a PLL, and
analog and mixed-signal blocks to support a high-speed serial interface. These are all designed to operate at a $V_{dd}$ of 0.9 V, chosen to provide simultaneous performance benefits and active power savings.

This library can be described as foundry-friendly because all key density-limiting ground rules of the technology are compatible with those of bulk technologies offered by foundries. Thus, migration of bulk designs does not necessitate wholesale physical layout changes. The performance and leakage characteristics of the SOI technology transistors are similar to or better than those of the bulk transistors offered by foundries for 90-nm technology. Thus, the SOI digital library offers an attractive migration path for designers needing a performance and power enhancement, but who also seek to avoid the costs and risks of not-yet-stable 90-nm technology. The schematic match between floating-body SOI transistors with bulk transistors makes it possible to use the same netlist to run circuit simulations in bulk and SOI technologies.

Designers may anticipate difficulties when migrating a design from bulk to SOI as a result of SOI-specific effects, such as the floating-body effect, local heating, and parasitic bipolar current. The higher leakage levels of an aggressively scaled technology are also an issue. However, in the experience of the IBM team that designed the SOI digital library, some problems were easily prevented, and enough design options were available to solve the others. For example, local heating in circuits with substantial dc current was found to be unimportant, because the physical layout of the starting-point bulk design provided ample thermal conductance. Transistors with high threshold voltage are available and were used to control leakage. Body-contacted devices were used in analog circuits for which transistor matching and output conductance are critical. Latches must always be handled with care in SOI because floating-body voltages make transistor strength drift, causing problems in weak feedback-style latches where keepers and inverters compete for control of the same node. Clocked feedback, in which the clock turns off the keeper during the write operation, is preferred and was used where possible to eliminate keeper-inverter competition and to allow large keeper devices to protect against noise. In an SOI technology, noise must be treated with special care because the parasitic bipolar current is an extra noise source.

During migration of the digital library design from bulk to SOI, the IBM design team found that much of the physical layout was reusable, which greatly simplified the process. **Figure 15** shows the physical layout of the chip. Almost all placement and wiring was reused. The thick-oxide transistors of the SOI technology were used to maintain the library's 2.5-V I/O. Though ESD diodes in SOI consume more chip area than their bulk counterparts, this discrepancy was handled without changing the I/O pin number or configuration.

**Low active and standby power SIMD unit**

Low-power DSP is a key component for wireless/wired SoCs. Communication DSP applications typically allow a high number of parallel single-instruction multiple-data (SIMD) operations without loss of computational efficiency. This in turn allows power savings through the use of many parallel gates running at moderate frequencies. SIMD width is often limited by the read and write port pressure on the register files supporting the SIMD pipelines. Optimal design points require a balance between high-frequency design and SIMD pipeline width to support high throughput at low power. The use of low-threshold voltage, thin-oxide, and short-channel SOI devices can allow very high frequency at minimum active power, but at the expense of increased subthreshold leakage power. These high-performance devices—combined with custom circuit techniques using low-power latches [42], data transition barriers, and tuned circuits—are ideal for maintaining high performance at low active power. While these technology choices are ideal for high-performance digital and analog circuits, they present a challenge for low-standby-power applications for large SIMD widths. For wireless applications, controlling standby power is a key issue. The most promising approach to solving this important issue is the use of high-$V_t$, low-leakage FETs as header or footer power switches [43, 44] that dynamically alter the pipeline width to trade
0.12-μm technology, the SOI FETs benefit from more aggressive engineering and have higher leakage than the bulk FETs. Nevertheless, these results demonstrate that outstanding active power can be achieved using high-performance SOI technology, and these basic SIMD building blocks can be used to construct an ultralow-power high-throughput DSP for wireless communication applications.

Isolation on SOI

The integration of large digital cores with sensitive analog, RF, and high-speed functions will require improved isolation. SOI technology provides a 10× lower coupling capacitance to the substrate and offers the opportunity to use higher-resistivity substrates (HRS). For example, a 10-mA SOI n-FET exhibits, at 5 GHz, a 37-kΩ isolation between the drain and the substrate. For a bulk device having the same current density, the isolation impedance would be 10× lower. In the discussion of sources of losses in Section 3, we saw that the Si substrate behaves like a high-pass first-order filter with a cutoff frequency given by

\[ f_{\text{cut}} = \frac{1}{2\pi R S C} \]

S-parameter measurements were performed on wafers from bulk and SOI technologies to quantify the enhanced isolation of SOI. The measured structure consists of two 50-μm × 50-μm n+ diffusions separated by 50 μm. SOI wafers with 12-Ω-cm and 100-Ω-cm substrates respectively showed isolation improvement of 13.6 dB and 22.4 dB compared with a bulk wafer with a 2-Ω-cm substrate. This demonstrates that significant isolation improvement can be achieved between large circuit blocks by using HRS SOI from dc to the GHz frequency range.

Low-power SOI technology also has the advantage of generating less noise. In the previous section, we saw an example of a digital circuit exhibiting 2× less power consumption on SOI and, therefore, 2× less digital noise.

5. Summary and conclusion

As is well known, SOI CMOS—as a result of its low diffusion capacitance—has inherent advantages as a low-power technology. But there are numerous other ways, less well known, to take advantage of SOI CMOS to enhance power savings and performance, and to do it with minimal design difficulties beyond standard bulk design. As shown in this paper, one can take advantage of the high performance of SOI to integrate high-speed communications circuitry with digital elements in a SoC with low overall system power and cost. Advantage can be taken of the functional equivalence of SOI and bulk CMOS technology to use existing proven power-management schemes. The near match between bulk and SOI physical layout and device behavior can be used to facilitate migration of bulk designs and libraries.
preserving their power and performance advantages. In short, efficient integration of wired and wireless transceivers with a microprocessor—made possible by low-power, high-performance HRS SOI technology—is fundamental to creating the always-connected and on-demand computing world that defines our future.

Acknowledgments
We acknowledge the contributions of our colleagues at the Advanced Semiconductor Technology Center, and the support of S. Chaloux, G. Shahidi, B. Davari, M. Dupuisquier, D. Friedman, M. Soyuer, and M. O Brysky.

References


Received December 2, 2002; accepted for publication March 27, 2003.

IBM J. RES. & DEV. VOL. 47 NO. 5/6 SEPTEMBER/NOVEMBER 2003

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Jean-Olivier Plouchart  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (plouchart@us.ibm.com). Dr. Plouchart received a Ph.D. degree in electrical engineering from Paris VI University in 1994. He spent ten months working at Alcatel Telspace on the design of MESFET GaAs MMICs for satellite telecommunications while pursuing a master's degree. From 1989 to 1990, he was a scientist/consultant at the ETCA as part of his military service. From 1990 to 1994 he was with the French Telecom Laboratory (CNET) as a Ph.D. student, developing the GaAs HBT technological process and designing HBT MMICs for DCS 1800 radio communications and high-bit-rate optical communications. In 1994 he joined the University of Michigan as a Research Fellow, designing high-speed circuits using a 100-GHz InP HBT process. Dr. Plouchart joined the IBM Thomas J. Watson Research Center as a research scientist in 1996; his work involved the design of SiGe BC莫斯 and CMOS RF/HIC circuits for wireless LAN applications. He currently manages the development of SOI technology for low-power and high-speed SoC applications at SRDC. Dr. Plouchart holds one U.S. patent. He was a coauthor of a Best Student Paper at RFIC, and has more than 40 publications in journals and conferences in the field of RF and high-speed design.

Noah Zamdmer  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (noahz@us.ibm.com). Dr. Zamdmer received a Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology in 1999. He has been working in SOI technology development at the IBM Microelectronics SRDC since 1999. Dr. Zamdmer's research interests include high-frequency device characterization and modeling, and the optimization of SOI CMOS technology for SoC applications.

Jonghae Kim  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (jonghae@us.ibm.com). Dr. Kim received a Ph.D. degree in electrical and computer engineering from the University of Minnesota at Minneapolis in 2001. From 1986 to 1995, he worked on the research and development of wireless and cellular systems at Samsung Electronics, Korea, where he worked on wireless transceiver design. While at Samsung Electronics, he developed many models of Samsung cellular systems which were based on AMPS, NMT, TDMA, and CDMA. From 1996 to 2001, he worked as a research scientist at the University of Minnesota, studying an SoC solution for ISM band wireless transceivers via CMOS. Dr. Kim received the first prize award from the 2000 SRC Copper Design Challenge. In 2001, he joined the Low-Power and RF Technology group of the IBM Microelectronics SRDC. He is currently working on SOI CMOS low-power and RF technology research activities. Dr. Kim's expertise and interests are in high-performance SoI passive devices, passive model development, low-power and RF circuit design, and SoC.

Melanie Sherony  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (sheronmy@us.ibm.com). Dr. Sherony received a B.S. degree in electrical engineering from the Massachusetts Institute of Technology in 1993 and 1997, respectively. While at MIT she specialized in extreme submicron SOI MOSFET design. She joined the IBM Microelectronics Division SRDC in 1998 and initially worked on development and device design for the high-performance 0.18-µm SOI CMOS technology. Dr. Sherony is currently working on the development of SOI CMOS technology for low-power, RF, and foundry applications.

Yue Tan  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (yueta@us.ibm.com). Dr. Tan joined the Hong Kong University of Science and Technology (HKUST) as a Research Fellow in 1999 after receiving a Ph.D. degree in electrical and electronic engineering from Southeast University, China. He designed and implemented the first fully integrated SOI RF power amplifier while at HKUST. He also developed a high-voltage (100 V) CMOS technology for plasma display panel applications. In 2001, Dr. Tan joined the IBM Microelectronics Division. He began work on SOI RF devices and integrated circuit design at the IBM SRDC, Hopewell Junction, New York, as a means of supporting early SOI device, technology, and electronic design automation learning for foundry business and product. Dr. Tan is now applying his interest in advanced techniques for physical verification, deep-submicron electronic design automation integration, and semiconductor device modeling at the IBM Microelectronics Division facility in Hopewell Junction, in support of IBM leading-edge SOI foundry products. He has focused predominantly on SOI device and circuit design issues associated with the implementation of integrated RF power amplifiers for wireless applications, and has authored or coauthored more than 20 technical papers in that field.

Asit Ray  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (ray@us.ibm.com). Dr. Ray received a Ph.D. degree from the University of Southern California at Los Angeles in 1977. He then joined the Exploratory Silicon Technology Department of the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, as a Research Staff Member. From 1977 to 1990, he worked on low-temperature silicon processing in the area of plasma oxidation, plasma nitridation, high-pressure oxidation, nitridation, annealing of process-induced damage, and multilevel metal interconnection. From 1990 to 1992, he was on assignment at Sematech, Austin, Texas, to work on rapid thermal processing tool and process development for 0.35-µm CMOS technology. Dr. Ray subsequently joined the Advanced Logic and SRAM Technology Development group of the IBM Microelectronics SRDC in Hopewell Junction, New York. He was involved in the process development and integration of 0.35-0.12-µm CMOS technology. For the past three years he has been working on the development of low-power digital and RF applications using SOI CMOS technology. Dr. Ray has published more than 60 papers on silicon technology and has presented numerous papers at international conferences.

Mohamed Talbi  *IBM Microelectronics Division, Semiconductor Research and Development Center, Route 52, Hopewell Junction, New York 12533* (talbi@us.ibm.com). Mr. Talbi received B.S. and M.S. degrees in industrial engineering and computer sciences from the University of Massachusetts in 1980 and 1984, respectively. He joined IBM in 1984 and initially worked on electronic packaging, modeling, and design. Mr. Talbi is currently working on design kit development for SOI technologies, mask data preparation, and design services.
Stephen V. Kosonocky  IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (stevek@us.ibm.com). Dr. Kosonocky received B.S., M.S., and Ph.D. degrees from Rutgers University in 1986, 1991, and 1994, respectively. From 1986 to 1992 he was a Research Scientist at Siemens Corporate Research in Princeton, New Jersey, where he worked on digital and analog circuit design. From 1992 to 1993 he was with the Samsung Princeton Design Center, where he worked on mixed-signal video circuits. In 1994 he joined the IBM Thomas J. Watson Research Center, where he has worked on embedded DRAM, systems-on-a-chip, circuit, logic, and microarchitecture design of Daisy and Box VLIW processors. Since 2000 Dr. Kosonocky has led a team on low-power circuits, where he is interested in developing ultralow-power signal processing cores for communication systems and low-power embedded PowerPC processors.

Daniel R. Knebel  IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (knebel@us.ibm.com). Mr. Knebel received a B.S. degree in electrical engineering from Purdue University. He joined the IBM Data Systems Division in 1982, developing circuit families for high-performance server processors. He has since worked in a variety of VLSI research and development areas, including microprocessor design, design verification, reliability, diagnostics, and electronic design automation. Mr. Knebel is a Senior Engineer in the VLSI Systems Department at the IBM Thomas J. Watson Research Center. He is currently a member of a research team developing energy-reduction techniques for current and future processor products. Mr. Knebel has received various IBM Achievement Awards and publication awards and has several patents pending.

Suhwan Kim  IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (suhwan@us.ibm.com). Dr. Kim received a Ph.D. degree in electrical engineering and computer science from the University of Michigan at Ann Arbor in 2001. That same year he joined the IBM Thomas J. Watson Research Center, where he is currently a Research Staff Member. His research interests encompass high-performance and low-power circuit/microarchitecture for SoC and low-power design methodologies for high-performance VLSI signal processing. Dr. Kim received the 1991 Best Student Paper Award of the IEEE Korea Section and the First Prize in the VLSI Design Contest of the 2001 ACM/IEEE Design Automation Conference. He has served as Workshop Chair for the IEEE International SoC conference and as Technical Program Committee Member for the IEEE International SoC conference and the International Symposium on Low Power Electronics and Design.

Keith A. Jenkins  IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (jenkinsk@us.ibm.com). Dr. Jenkins received a Ph.D. degree in physics from Columbia University for experimental work in high-energy physics. He is a Research Staff Member at the IBM Thomas J. Watson Research Center, where he is a member of the Communications Technology Department. At the IBM Research Division, he has done research in a variety of device and circuit subjects, including high-frequency measurement techniques, electron beam circuit testing, radiation-device interactions, low-temperature electronics, and SOI technology. His current activities include evaluation of the performance of VLSI circuits and phase-locked loops and investigations into substrate coupling in mixed-signal and RF circuits. Dr. Jenkins is a Senior Member of the IEEE.

Michel M. Rivier  IBM Microelectronics Division, Essex Component Development Laboratory, 91105, Corbeil-Essonnes, France (rivierfe@fr.ibm.com). Dr. Rivier received a Ph.D. degree from Stanford University in 1982. He began his career as a Research Staff Member at the Bendix Aerospace Technology Center, where he did SOI process and device work using laser recrystallization. In 1985 he joined IBM France, where he held several positions in the Essex Component Development Laboratory. Dr. Rivier is currently a Senior Engineer in charge of various analog and mixed-signal chips and core development for internal and external customers, based on a broad range of technologies such as CMOS, SOI, BiCMOS, and SiGe.