Latency and Power Measurements on a 64-kb Hybrid Josephson-CMOS Memory

Q. Liu, K. Fujiwara, X. Meng, S. R. Whiteley, T. Van Duzer, N. Yoshikawa, Y. Thakahashi, T. Hikida, and N. Kawai

Abstract—A 64-kbit sub-nanosecond Josephson-CMOS hybrid RAM memory is being developed with hybrid high-speed interface circuits. The hybrid memory is designed and fabricated using a commercially available $0.18\ \mu\text{m}$ CMOS process and NEC-SRL's $2.5\ \text{kA}/\text{cm}^2$ Nb process for Josephson junctions. The memory bit-line output signals are detected by ultra-low power, high-speed Josephson devices. The most challenging part of the memory system is the input amplifier; the performance of this amplifier is optimized by minimizing its parasitic capacitance loading. Functionality tests were reported at the ASC 2004, high-speed measurements are the focus of this paper. Most of the power dissipation in the memory system occurs in the input interface circuits. The power dissipation and delay for the interface circuit are reported. The delay time and power dissipation of the CMOS part (including drivers, decoders, and cells), which contributes most to the total access time of the hybrid memory, are measured. The relation of measurements to simulations will be presented. Plans for future designs to reduce power dissipation and latency are described.

Index Terms—Access time, high-speed measurement, hybrid memory, interface circuit.

I. INTRODUCTION

JOSEPHSON-CMOS hybrid random-access memories (RAMs) have the potential to remove the memory bottleneck faced by Josephson junction (JJ) digital technology [1]–[3]. The main idea is to use high-density charge-storage MOS memory cells and access them by high-speed ultra-low-power superconductive devices, which takes advantage of the best features of each technology. Fig. 1 shows the system block diagram. This approach was proposed years ago [3] and some preliminary simulations and measurements have been reported [4]–[6], which verifies the feasibility of this idea. In this paper, high-speed measurements of the 4 K CMOS memory part and the hybrid interface circuit will be presented, both delay times and power consumptions will

Fig. 1. System block diagram of the hybrid memory with a standard CMOS 3-T memory cell schematic in the inset.

be reported, and the relation between simulation results and measurement results will be discussed.

II. 4.2 K CMOS MODELING

There are two key issues in this approach. The first one is the low-temperature (4 K) operation of commercial CMOS circuits. We have reported [7] that, for several sub-micron commercial CMOS processes, both individual devices and circuits work better at 4 K than at room temperature. Details of 4 K CMOS operation can be found in [7]. But in summary, the speed of a standard digital circuit will be improved by 40% to 50% from room temperature to 4 K, with the power consumption reduced by up to 30%, depending on the operation of the circuit. In this paper, a commercial $0.18\ \mu\text{m}$ dual-well dual-gate CMOS process was used to fabricate the CMOS memory chips. Based on experimental results, 4 K operation of these $0.18\ \mu\text{m}$ CMOS chips compares well to what has been previously been reported for the $0.25\ \mu\text{m}$ process, without having any new physical phenomena that cannot be explained by the modified room-temperature BSIM3 model.

III. HYBRID INTERFACE CIRCUIT

The other key issue is the interface circuit, which interfaces between millivolt-level superconductor signals and volt-level CMOS signals, in a fast and power-efficient manner. The so-called Suzuki stack [8], which has been studied intensively, is a good candidate for the first stage of the interface circuit. The delay time is about 20 ps for a Suzuki stack with a 40 mV output, and operation up to several gigahertz is possible. There

Manuscript received August 29, 2006. The work at Berkeley was supported by the Office of Naval Research under Grant No. N00014-03-1-0065. The work at Yokohama was supported by VLSI Design and Education Ctr. (VDEC), the University of Tokyo, in collaboration with Cadence Design Sys., Inc. and Synopsys, Inc.

Q. Liu, K. Fujiwara, X. Meng, S. R. Whiteley, and T. Van Duzer, are with the Department of EECS and Electronics Research Lab (ERL), University of California, Berkeley, CA 94720 USA (e-mail: vanduzer@eeecs.berkeley.edu).

N. Yoshikawa, Y. Thakahashi, T. Hikida, and N. Kawai are with the Department of Electrical and Computer Engineering, Yokohama National University, Hodogaya, Yokohama, Japan.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASC.2007.898698

Fig. 2. Schematic of fast hybrid interface circuit proposed by Ghoshal *et al.* [3]. A clocked input pulse current feeding the following dual stack.

are several candidates for the second-stage circuit, which amplifies 40 mV to about 1 V. Although pure CMOS differential amplifiers accomplish the amplification process quickly, they consume disadvantageously large power, and there is no way to decrease the power without sacrificing delay time..

The hybrid interface [6] is the best choice in terms of power. Fig. 2 shows the whole circuit with a 40-mV-output Suzuki stack. The load on the NMOS device M_1 which receives a 40 mV signal from the Suzuki stack is a 400-JJ series array having a critical current of $400 \mu\text{A}$ in our present design. The device M_1 is biased so that the load current is 80% of the critical current of the JJs in the load. With our presently available Josephson niobium technology, the total spread of I_c in the array is just a few percent [9], so all junctions are initially in the zero-voltage state. The 40 mV coming from the Suzuki stack increases the transconductance of M_1 , thus driving more current into the junction array, and switching all 400 junctions. The delay time depends on the discharging process of the output node that occurs when the junctions go into the voltage state, so the output-to-ground capacitance is very critical. Several approaches have been reported [6] to decrease the parasitic capacitance and it is believed that the capacitance is lowered to 80 fF for a $6.5 \text{ kA}/\text{cm}^2$ Nb process. A delay of 100 ps is calculated using WRSPICE simulation [10] based on this capacitance value. In this paper, the JJ chips were fabricated using a $2.5 \text{ kA}/\text{cm}^2$ Nb process; due to the larger size of the junctions, the capacitance of the array is around 130 fF and the delay of the second stage is simulated to be 170 ps.

IV. SAMPLE PREPARATION

The chips for the hybrid memory were fabricated using a commercially available standard $0.18 \mu\text{m}$ CMOS process and NEC-SRL's $2.5 \text{ kA}/\text{cm}^2$ Nb JJ process [11]. The physical sizes are $2.4 \text{ mm} \times 2.4 \text{ mm}$ for the CMOS chip, where a 64-kb random-access memory with decoders and address buffers occupies about $1.2 \text{ mm} \times 1.2 \text{ mm}$ area, and $5 \text{ mm} \times 5 \text{ mm}$ for the JJ chip. We did not have solder-bump bonding available, so the CMOS chip was mounted face-up on top of the JJ chip and the interconnections were made with very short 50- μm -diameter aluminum wire bonds. Fig. 3 shows a picture of one piggy-backed chip set. The CMOS chip is thinned to about $200 \mu\text{m}$ before attaching it to the JJ chip to decrease the wire

Fig. 3. The photograph of the piggy-back chip. The CMOS chip is $2.4 \text{ mm} \times 2.4 \text{ mm}$ and the JJ chip is $5 \text{ mm} \times 5 \text{ mm}$. The cross point of the two perpendicular stripes is where the measured cell is located.

Fig. 4. Schematic of delay-measurement circuit. The signal is triggered when CLK is high and D_CLK is low. DUT (device under test) could be an interface circuit or CMOS memory.

bond length and concomitant inductive parasitics, which are harmful to circuit operation. Pad locations are arranged such that possible crosstalk will be minimized and wire lengths are minimized. After the wire-bonding, the piggy-back chip set is mounted on a BCP-2 cryoprobe, a wide-bandwidth probe manufactured by American Cryoprobe Inc. The probe was modified to accommodate the CMOS chip.

V. EXPERIMENTS

Simulations have shown that gigahertz operation for the hybrid memory is possible and sub-nanosecond access time is expected. However, it is difficult to measure such small delay time in traditional ways. In the semiconductor field, it is common to use ring-oscillators (a loop comprising an odd number of inverters), with the delay time of an individual inverter obtained from the oscillation frequency. This idea, however, does not apply to our memory and interface measurements. Ghoshal [12] proposed the hybrid circuit in Fig. 4 to measure directly small individual delays. The MOS devices M_1 and M_2 are designed such that the "ON" current is high enough to switch the 4-JJ arrays as well as the Suzuki stack in the DUT. The MOSFET

Fig. 5. Delay of second part of interface circuit. A 430 ps delay is measured, with $V_{DD} = CLK = 1.5$ V.

Fig. 6. The simulation result for the second part of the interface amplifier shows that a large delay (310 ps) is incurred in obtaining the necessary 0.7 V to drive the next stage.

M_3 is identical to M_1 and M_2 so that any parasitic effects will be compensated in the measurement. The delay between OUT_1 and OUT_2 is the delay of the interface circuit. The two cables from OUT_1 and OUT_2 to the oscilloscope must be exactly the same in length. With careful set up, the accuracy of this measurement is believed to be about 20 ps.

Fig. 5 shows the delay-measurement result for the second stage of the hybrid interface circuit (part 2 in Fig. 2). From the picture, one can read a 430-ps delay. This value is larger than the 170 ps found from simulation. The possible reasons follow. The delay time simulated previously was based on the difference between the half-level points of the input and output. But half of the voltage drop of the interface circuit, 0.5 V, is insufficient to drive the following PMOS with high enough current to switch the 4-JJ output array. Rather, 0.7 V is required.

Fig. 6 shows the simulation curves for the second stage of the interface amplifier where it can be seen that the delay would be 310 ps to obtain a 0.7 V drop. Thus 140 ps of the difference (260 ps) between the measurement and simulation is accounted for. Also, unidentified parasitics from the bonding wires or pad connections may increase the delay time.

One can expect that a high-current-density JJ process will decrease the delay time of the interface amplifier due to the smaller junction capacitances. Based on simulation, we believe

Fig. 7. Memory delay measurement waveforms including input CMOS driver, decoder, memory cell, and bit-line JJ readout. About 500 ps delay time is measured, with $V_{DD} = CLK = 1.5$ V.

that $10\text{ kA}/\text{cm}^2$ JJ process will decrease the delay time to less than 150 ps. Furthermore, we can take advantage of the sharp subthreshold property of 4 K MOSFETs and make the PMOS M_3 V_{DD} 0.3 V higher than the interface V_{DD} . When the interface circuit is off, this 0.3 V drop (V_{SG}) is not large enough to turn the following PMOS on, causing a problem, however, when the interface circuit switches, the 0.5 V drop in 400 JJ is equivalent to a 0.8 V drop for the following PMOS. By doing this, the delay time would be less than 100 ps for higher JJ current densities.

We have made successful functionality measurements of the complete interface amplifier. However, although simulations indicate high-speed measurements of the complete circuit should be successful, there is some experimental problem, apparently with the connection of parts 1 and 2 (Fig. 2). This suggests a parasitic problem, probably with the use of wire bonding, which should not be a problem if we use solder-bump bonding.

Fig. 7 shows the result of the 4 K CMOS part of memory delay test. The same measurement circuit (Fig. 3) was used with the current flowing out of M_2 charging up the input of the CMOS memory to V_{DD} , causing a “0” to “1” transition. This transition triggered by M_2 passes through the address buffer (inverter chain), decoders, and the memory cell; a bit line current switches the 4 JJ array if the cell stores a “1”. A 500 ps delay time can be read from the picture, which is somewhat larger than the simulation result of 400 ps. Reasons for the difference are still under investigation.

This delay time is from the CMOS address buffer input to the bit-line current output. Since the critical current of the 4-JJ output array is designed to be $200\text{ }\mu\text{A}$, which is almost the same value as the input current level of superconductor sensors which read out the bit-line current [6], we can conclude that this delay almost represents the real situation of the hybrid memory system. Adding the interface, memory, and sensor delays, the total access time of the current 64 kb memory is about 900 ps ($0.25\text{ }\mu\text{m}$ CMOS and $2.5\text{ kA}/\text{cm}^2$ JJ). We believe that with some improvement of the circuit structure (such as bump

Fig. 8. Power dissipation scales up with memory size.

bonding to replace the wire-bonding), 500 ps access time is possible for a 64 kb hybrid memory made of 0.18 μm CMOS and 10 kA/cm^2 JJ processes. Still further improvements could be achieved with a 20 kA/cm^2 JJ process and a 90 nm CMOS process specially designed for 4 K operation.

VI. POWER CONSUMPTION

The dynamic power of the CMOS parts of the memory depend on frequency, as described by $P = 1/2 CV^2 f$. The measured CMOS memory dynamic power for the reading process is 0.8 mW at 1 GHz, and for the writing process is 1.6 mW. The calculated power is around 0.7 mW for reading and 1.4 mW for writing. Pad and wire capacitances are believed to be the main contributors to the difference between calculated values and measured ones.

Power consumption of this memory occurs mostly in the interface amplifiers. It is 0.3 mW for an individual hybrid interface circuit, measured at 1 GHz. The total power of a hybrid memory depends strongly on how many interface circuits are used. Power consumption scales down with process scaling because both supply voltage and capacitances scale with CMOS process scaling. We expect less power consumption as we use more advanced technologies [13]. For larger sizes of memories, power consumption goes up very slowly as the memory size increases [13]. Fig. 8 shows the calculation of how power is related to the memory size.

VII. CONCLUSIONS

Recent progress of the hybrid memory project has been reported. Components of the access time and the power consump-

tion of a 64-kb hybrid memory were presented. The measured delays are somewhat larger than simulated ones, and the possible reasons are discussed, as well as possible improvements. We believe that the total access time will be around 500 ps with a 10 kA/cm^2 JJ process and 0.18 μm CMOS. The measured power consumption fit well the calculated ones. Both access time and power consumption could be further improved if more advanced technologies are used.

ACKNOWLEDGMENT

Q. Liu thanks Jim Wieser of National Semiconductor for his assistance on the CMOS chips. N. Yoshikawa thanks the process group at Superconducting Research Laboratory in Tsukuba, Japan for the foundry service of the Nb Josephson ICs.

REFERENCES

- [1] S. V. Polonsky, A. F. Kirichenko, V. K. Semenov, and K. K. Likharev, "Rapid single flux quantum random access memory," *IEEE Trans. Appl. Supercond.*, vol. 5, pp. 3000–3005, June 1995.
- [2] S. Nasagawa, H. Numata, Y. Hashimoto, and S. Tahara, "High-frequency clock operation of Josephson 256-word 16 bit RAMs," *IEEE Trans. Appl. Supercond.*, vol. 9, pp. 3708–3713, June 1999.
- [3] U. Ghoshal, H. Kroger, and T. Van Duzer, "Superconductor-semiconductor memories," *IEEE Trans. Appl. Supercond.*, vol. 3, pp. 2315–2318, March 1993.
- [4] F. J. Feng, X. Meng, S. R. Whiteley, T. Van Duzer, K. Fujiwara, H. Miyakawa, and N. Yoshikawa, "Josephson-CMOS hybrid memory with ultra-high-speed interface circuit," *IEEE Trans. Appl. Supercond.*, vol. 13, pp. 467–470, June 2003.
- [5] T. Van Duzer, Q. Liu, X. Meng, S. R. Whiteley, and N. Yoshikawa, "High-speed interface amplifiers for SFQ-to-CMOS signal conversion," in *Extended Abstracts, International Superconductor Electronics Conference, ISEC'03*, Sydney, Australia, June 2003.
- [6] Q. Liu, T. Van Duzer, X. Meng, S. R. Whiteley, K. Fujiwara, T. Tomida, K. Tokuda, and N. Yoshikawa, "Simulation and measurements on a 64-kbit hybrid Josephson-CMOS memory," *IEEE Trans. Appl. Supercond.*, vol. 15, pp. 415–418, June 2005.
- [7] N. Yoshikawa, T. Tomida, M. Tokuda, Q. Liu, X. Meng, S. R. Whiteley, and T. Van Duzer, "Characterization of 4 K CMOS devices and circuits for hybrid Josephson-CMOS systems," *IEEE Trans. Appl. Supercond.*, vol. 15, pp. 267–271, June 2005.
- [8] H. Suzuki, A. Inoue, T. Imamura, and S. Hasuo, "A Josephson driver to interface Josephson junctions to semiconductor transistors," in *Dig. Tech. Papers, Int. Electron Devices Meeting*, December 1988, pp. 290–293.
- [9] X. Meng, L. Zheng, A. Wong, and T. Van Duzer, "Micron and submicron Nb/Al-AlOx/Nb tunnel junctions with high critical current densities," *IEEE Trans. Appl. Supercond.*, vol. 11, pp. 365–368, March 2001.
- [10] [Online]. Available: <http://www.wrcad.com/wrspice.html>
- [11] S. Nagasawa, Y. Hashimoto, H. Numata, and S. Tahara, "A 380 ps, 9.5 mW Josephson 4-Kbit RAM operated at a high bit yield," *IEEE Trans. Appl. Supercond.*, vol. 5, pp. 2447–2452, June 1995.
- [12] U. Ghoshal, "Ph. D. dissertation, University of California, Berkeley, CA, 1995.
- [13] Q. Liu, T. Van Duzer, K. Fujiwara, and N. Yoshikawa, "Hybrid Josephson-CMOS memory in advanced technologies and larger sizes," in *7th European Conference on Applied Superconductivity, Journal of Physics: Conference Series 43 (2006)*, 2005, pp. 1171–1174.