

## Hybrid Josephson-CMOS Memory in Advanced Technologies and Larger Sizes

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2006 J. Phys.: Conf. Ser. 43 1171

(<http://iopscience.iop.org/1742-6596/43/1/285>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 38.107.179.214

The article was downloaded on 15/02/2012 at 01:54

Please note that [terms and conditions apply](#).

# Hybrid Josephson-CMOS Memory in Advanced Technologies and Larger Sizes

Q. Liu<sup>1</sup>, T. Van Duzer<sup>1</sup>, K. Fujiwara<sup>1,2</sup>, and N. Yoshikawa<sup>2</sup>

<sup>1</sup> Dept. of Electrical Engineering and Computer Science, University of California, Berkeley, CA94720, USA

<sup>2</sup> Department of Electrical and Computer Engineering, Yokohama National University, Hodogaya, Yokohama, Japan

Email : vanduzer@eecs.berkeley.edu

**Abstract.** Recent progress on demonstrating components of the 64 kb Josephson-CMOS hybrid memory has encouraged exploration of the advancement possible with use of advanced technologies for both the Josephson and CMOS parts of the memory, as well as considerations of the effect of memory size on access time and power dissipation. The simulations to be reported depend on the use of an approximate model for 90 nm CMOS at 4 K. This model is an extension of the one we developed for 0.25  $\mu\text{m}$  CMOS and have already verified. For the Josephson parts, we have chosen 20 kA/cm<sup>2</sup> technology, which was recently demonstrated. The calculations show that power dissipation and access time increase rather slowly with increasing size of the memory.

## 1. Introduction

Josephson-junction-based RSFQ logic circuits comprise the only technology with demonstrated capability of an order of magnitude increase in speed over CMOS logic while having orders of magnitude lower power dissipation. Random access memory is needed to complement the RSFQ logic. However, lack of large-size, dense, and sufficiently fast memory has been a long-standing problem in superconducting digital technology, especially for high-end computing applications. Some demonstrated memory systems using only Josephson junction technology were both insufficiently dense and too small for high-end applications. [1, 2] The Josephson-CMOS hybrid memory (Fig. 1) was proposed in 1993 to try to solve the problem. [3] It has the potential to remove the memory bottleneck faced by Josephson digital technology. The main idea is to use high-density charge-storage MOS memory cells and access them by high-speed superconductive devices. This takes advantage of the best features of each technology. In this paper, we review both current and more advanced CMOS and Josephson technologies for hybrid memory. Simulations will be presented for various sizes of hybrid memory using different technologies.

## 2. A 64 kb hybrid memory system using current technologies

Since hybrid memories have to work in superconducting circuit environments (4 K for Nb technology), 4 K sub-micron CMOS operations have been extensively explored and a well-fitting BSIM3 model for standard commercial 0.25  $\mu\text{m}$  process CMOS has been established. According to the 4 K model, operating sub-micron CMOS devices at 4 K will increase memory circuit speed

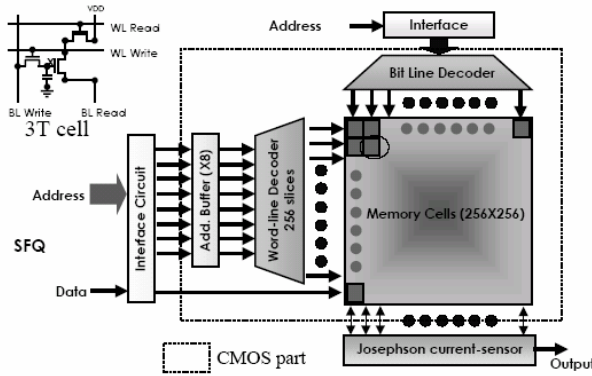


Fig. 1 Hybrid memory system

are pure CMOS or BJT amplifiers and a Josephson-CMOS hybrid amplifier. The semiconductor amplifiers consume much more power than the system can afford, even though they are simple to design and fabricate. The hybrid one, on the other hand, consumes less power and runs faster than the pure CMOS one by paying in the complexity of a hybrid technology. The hybrid amplifier shown on Fig. 2 has been chosen. The preamplifier (left one) uses a dual series array of junctions. We have added a preceding 4JL stage to convert an SFQ pulse (usually 1-2 ps wide with amplitude less than 1 mV) to a 200  $\mu$ A current feeding the dual series array. The output of the preamplifier is about 40 mV with duration equal to that of the clock  $V_{dd}$  duration and its delay is about 20 ps [5]. The next part consists of three MOSFETs and a 400 Josephson junction array load. The bottom transistor is biased so that the load current is 80% of the critical current of the Josephson junctions in the load. With current Josephson niobium technology, the total spread of  $I_c$  in the array is just a few percent [6]; so all junctions are initially in the zero-voltage state. When the 40 mV pulse drives the gate of the bottom transistor, which acts as a voltage-controlled current source to increase the load current above the critical current of junctions, the junctions will switch and generate a swing of about 1.2 V at the output

Extensive simulations of the amplifier have been done and reported in [5], taking account of parasitic inductances and capacitances in the load circuit containing the 400-junction array. The switching delay with achievable parasitics is 80 ps. Added to  $\sim$ 20 ps for the preamplifier, this gives a total amplifier delay of 100 ps. Power consumption of the amplifier due to the DC bias is higher than desired. The power is mostly static and for current design is 0.6 mW. The dynamic power only contributes 10  $\mu$ W more to the total power if the frequency is 1 GHz.

compared with 300 K operation as well as allow operation at low voltage, resulting in reduced power dissipation. Another important benefit of operating at low temperature is that the leakage currents are substantially suppressed so that there is no need to refresh the memory cells. [4]

2.1. Interface circuit

The interface circuit plays a key role in hybrid memory. It has to convert sub-milivolt picosecond SFQ pulses into volt-level CMOS digital signals as quickly as possible, with affordable power consumption. Candidates

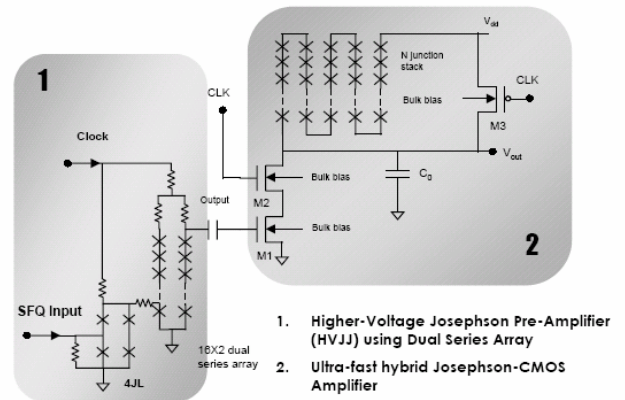


Fig. 2 Hybrid interface circuit

2.2. Address buffer, decoder and memory cell (CMOS part)

Circuits in this part employ pure CMOS technology. Address buffers are typical inverter buffers to minimize delay time. And decoders are built up with static CMOS logic in order to make the circuits robust and consume lower power. Approaches to designing static CMOS decoders vary depending on the system requirement. In the hybrid memory system, the delay time of the decoder is the main part of the total access time; therefore we focused in this work on how to minimize delay. The 3-T DRAM cell was chosen as the hybrid memory cell to get non-destructive readout and sufficient density. Due

to the low-temperature CMOS properties, the dynamic RAM behaves like a static one in that the need of refresh circuits as used in typical 300 K semiconductor memory is eliminated. After choosing appropriate logic and making a careful design, the simulation delay for the address buffer and decoder with proper load is about 400 ps for 0.25  $\mu\text{m}$  technology. Since static CMOS logic is used, the power consumption comes almost all from the dynamic power, which is proportional to  $CV^2f$ , less than 1.5 mW for 1 GHz operation of an 8-input decoder.

### 2.3. SFQ sensor circuit

A 4JL based SFQ sensor was reported in [6]. Simulation results indicate that ultra-fast and low-power detection can be expected by using this bit-line detector and its bias margin is about  $\pm 18\%$ . Pure CMOS memory suffers from threshold drop, so a sensor amplifier with gain somewhat greater than unity is needed to amplify the signal to full swing. Therefore, an extra delay and a large amount of power dissipation are introduced. In contrast, superconducting sensor consumes only 18  $\mu\text{W}$ .

### 2.4. Operation of memory and power dissipation

The total power consumption depends on the operation mode and the read/write scheme. For high-end computation application, i.e., an SFQ computer, the memory will receive data from an SFQ CPU, so the data width depends on the word size of the SFQ CPU. For a 32-bit CPU system, the CPU will write and read in the hybrid memory complete 32-bit words. For the 64 kb memory, the reading process needs 11 input amplifiers plus 32 sensors, and the total power for a 32-bit read-out process is about 10 mW. The writing process, on the other hand, requires 11 input hybrid amplifiers for the address decoding and another 32 for the data to be written. Taking all parts into account, the writing power is about 28 mW. All power estimates above assume 1 GHz operation with  $V_{\text{DD}}$  of CMOS being 1.5 V.

The simulation of the hybrid memory shows that the total access time is about 0.5 ns for the assumed technologies (6.5  $\text{kA}/\text{cm}^2$  Nb process and 0.25  $\mu\text{m}$  standard CMOS process). Low-speed operation of each part and the system-level functionality have been demonstrated [7]. High-speed measurements (above 1 GHz) are now in progress.

## 3. More advanced technologies for 64 kb hybrid memory

With the continued development of both semiconductor and superconductor technologies, more advanced processes are expected to be used in the hybrid memory in the near future. A 20  $\text{kA}/\text{cm}^2$  Nb process has been demonstrated in the laboratory [8] and 90 nm CMOS processes are now industrial standard. Based on the scaling rules of CMOS and assuming that the low-temperature improvement of CMOS does not change much with scaling, we can calculate roughly the delay and power consumption for 64 kb hybrid memory using advanced technologies. Since the delay comes mostly from the CMOS part and most power dissipation is the interface static power (proportional to  $V_{\text{DD}}$ ), one can conclude that the upgrade of the CMOS process will affect the performance of hybrid memory more than the upgrade of Josephson process. Scaling rule calculations suggest that the access time for 90 nm CMOS will be 240 ps while power consumption decreases little because the interface static power dominates.

We must point out that the current and advanced CMOS processes referred to above are just standard CMOS processes designed for room-temperature operation, not for 4 K. There is plenty of room to improve the design of processes if working at 4 K is targeted. The most promising goals of the special design are reductions of  $V_{\text{TH}}$  and  $V_{\text{DD}}$ . In room temperature operation, one cannot make the threshold voltage lower than 0.3 V or  $V_{\text{DD}}$  lower than four times  $V_{\text{TH}}$ . The subthreshold current is too large and therefore the leakage current is too large if  $V_{\text{TH}}$  is less than 0.3 V, and the device does not work ideally if  $V_{\text{DD}}$  is less than four times  $V_{\text{TH}}$ . This sets the limit of the CMOS scaling rule for  $V_{\text{TH}}$  and  $V_{\text{DD}}$ . That is part of the reason the power is now becoming more and more important in semiconductor industry. The main performance advantage of 90 nm devices specially designed for 4 K operation, when compared to conventional 90 nm CMOS operated at 4 K, lies in scaling of the operating voltages, and thereby reduction of power dissipation even at higher frequency operation. It is possible to design threshold voltages of  $\sim 30 \text{ mV} \pm 5 \text{ mV}$  and operate circuits reliably with much lower  $V_{\text{DD}}$ . In

this way, one can achieve significant reduction in power dissipation. It also makes conversion of signals to and from SFQ circuits much easier.

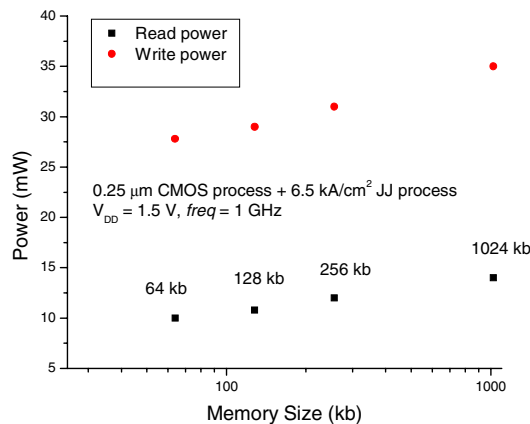


Fig. 3 Power dissipation with memory size scales up

Applying the design techniques called “logic effort” and “path effort”, one can easily find that the delay of the decoder increases only 70 % and its power consumption triples, compared with the decoder in a 64 kb memory. As to other parts of the hybrid memory, the number of interface amplifiers increases by four, which affects the total power consumption. In a word, the increase of memory size means a slow increase of delay and power consumption of CMOS circuits and the number of input amplifiers. The total access time for a 1 Mb hybrid memory using a 0.25  $\mu\text{m}$  CMOS process and 6.5  $\text{kA}/\text{cm}^2$  Nb process is about 0.8 ns. And the total power consumption for read and write process are 14 mW and 35 mW, respectively. Power dissipation for different sizes of memory is shown in Fig. 3. If a 90 nm CMOS process and a 20  $\text{kA}/\text{cm}^2$  are used, the total access time will be decreased to about 0.35 ns. Furthermore, the power can be significant reduced by using specially designed low-temperature CMOS process.

## 5. Discussion and conclusion

Both simulations and preliminary measurements have proven the hybrid memory to be a strong candidate for RSFQ logic systems. With more advanced technologies, both performance and power dissipation will be improved. Increasing the size of the memory increases delay and power, however, only slowly. Therefore, larger sizes of the hybrid memory with fast access time and low power dissipation are possible in the future.

## References

- [1] S. Nagasawa et al., *IEEE Trans. Appl. Superconduct.*, vol. 9, pp. 3708–3713, June 1999.
- [2] A.F. Kirichenko et al., *IEEE Trans. Appl. Superconduct.*, vol. 11, pp. 537-540, March 2001
- [3] U. Ghoshal et al., *IEEE Trans. Appl. Superconduct.*, vol. 3, pp. 2315-2318, March 1993
- [4] N. Yoshikawa et al., *IEEE Trans. Appl. Superconduct.*, vol. 15, pp. 267-271, June 2005
- [5] T. Van Duzer, ISEC 2003, Sydney, Australia, July 2003
- [6] H. Nakagawa et al., *Jpn. J. Appl. Phys.* **21** L198-L200, 1982
- [7] Q. Liu et al., *IEEE Trans. Appl. Superconduct.*, vol. 15, pp. 415-418, June 2005
- [8] L.A. Abelson et al., *Proc. IEEE*, vol. 92, pp. 1517-1533, October 2004

## 4. Larger size memories up to 1 Mb

The 64 kb memory is our first demonstration memory; however, it is not the ultimate goal for high-end computing. Larger size memories up to 1 Mb are needed. Even larger memory could be built up with 1 Mb banks using a multi-bank structure, as it is done in semiconductor industry.

The CMOS part of a 1 Mb memory could be laid out as 1024  $\times$  1024 square matrix, which requires a 10-input row decoder and a 5-input column decoder to access 32-bit words. The logic structure and style for decoders then change. And at the same time, the load capacitance for the decoder increases by a factor of four compared with 64 kb memory. However, this does not necessarily mean the delay and power increases by a factor of four, if carefully designed.