ELSEVIER

Physica B 280 (2000) 495–496

PHYSICA B

www.elsevier.com/locate/physb

COOL-1: the next step in RSFQ computer design

Konstantin Likharev[a,*], Mikhail Dorojevets[b], Paul Bunyk[b], Dmitry Zinoviev[a,1]

[a]Department of Physics and Astronomy, SUNY at Stony Brook, Stony Brook, NY 11794-3800, USA
[b]Department of Electrical and Computer Engineering, SUNY at Stony Brook, Stony Brook, NY 11794-2350, USA

Abstract

Recent progress in the design of a superconductor digital system for high-performance computing, based on the RSFQ logic family, is briefly outlined. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Josephson effect; RSFQ logic; High-performance computing

RSFQ logic circuits, based on the presentation of digital bits by single quanta of magnetic flux [1,2], can operate at clock frequencies beyond 100 GHz, consuming extremely low power (of the order of 10^{-18} J/bit, i.e. some 5 orders of magnitude lower than modern CMOS circuits). This small power dissipation allows very high chip packaging density, thus reducing interchip signal propagation delays and enabling the ultra-high operation speed to be sustained even in complex digital systems. At the same time, LTS RSFQ circuits are inherently much simpler to fabricate than CMOS circuits with comparable minimum feature size. For these reasons RSFQ has been selected as the core number-crunching technology in the large-scale hybrid technology multithreaded (HTMT) computer project [3] whose ultimate goal is the development of a petaflops-scale computing facility by or soon after year 2006.

Last year, our team released a draft design ("COOL-0") of the petaflops RSFQ subsystem [4], based on the assumption of 0.8-μm Nb-trilayer technology. The design shows that the peak performance of 1 petaflops may be provided by a very compact system of 4096 RSFQ processors ("SPELLs") with a total volume of about 0.5 m^3 and dissipated power (at 4.2 K) of only 250 W. However, the COOL-0 design also revealed a few challenges faced by the project. Most importantly, the total volume of the cryomemory ("CRAM") was low, only 1 MByte per SPELL, i.e. 4 GBytes for the whole RSFQ subsystem. This is a factor of 2096 smaller than the planned volume of the next (SRAM) layer of the HTMT memory hierarchy. This disparity puts an unreasonably high load on the cryo/room temperature interface: in COOL-0 it should have as many as 8 million 8-Gbps cables. Another challenge faced by COOL-0 was the number of contact pin on RSFQ dies which should be flip-chip mounted onto cryo-multichip modules ("CMCMs"). In particular, the chip housing functional units and interprocessor switching network ("PNET") had to have about 10000 pins. This is a very high toll indeed; the semiconductor electronics industry plans to approach this packaging level only by year 2010 or so.

Fortunately, the past few years have evidenced very substantial progress in the fabrication of Nb-trilayer Josephson junctions. Most significantly, deep-submicron junctions with critical current density up to 200 kA/cm^2 have been demonstrated [5,6]. In accordance with theoretical expectations, these junctions are intrinsically overdamped, and can be used in RSFQ circuits without external shunting. In addition, these junctions have high values of the $I_C R_N$ product (about 2.0 mV, i.e., very close to the theoretical maximum for niobium), and as a result may bring up speed of RSFQ circuits. For example, T flip-flops (i.e. digital frequency dividers) using these junctions have been demonstrated [6] to operate up to 770 GHz. Even more importantly, these self-shunted junctions have turned out to be quite reproducible, providing total on-chip spreads of I_c below 10%. The second important advance was the development [7] of vertically

* Corresponding author.
E-mail address: klikharev@sunysb.edu (K. Likharev)

[1] This work was supported in part by DARPA, NSA and NASA via Caltech, within the framework of the HTMT computer project.

0921-4526/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0921-4526(99)01845-1

stacked Josephson junctions with an external contact to the middle (common) electrode. These double junction stacks allow very natural implementation of compact RSFQ circuits.

Combined, these achievements should allow a drastic increase in RSFQ circuit density. For example, a typical RSFQ cell such as the SFQ pulse merger (i.e. the asynchronous OR gate [1]) using double-stack, 0.4-μm junctions can be laid out on an area of just 15 μm^2, i.e., below 100 lithographic squares. For comparison, in a standard current, 3.5-μm technology this cell occupies more that 3000 μm^2 (about 300 lithographic squares). Conservative estimates show that this new approach should allow RSFQ logic circuits equivalent in density to CMOS with 15 million transistors/cm^2, and 16 Mbit CRAM chips.

This advance changes the HTMT design situation significantly. First, CRAM finds a respectable place in the HTMT memory hierarchy, becoming (with 4 CRAM chips per SPELL) just 32 times smaller than SRAM. With an appropriate program execution model, this may decrease the requirements to the room temperature interface bandwidth. Second, it seem that now the complete SPELL may be laid out on a single chip, reducing the number of external contact pads.

We have accepted the deep-submicron, self-shunted JJ technology as the basis for our next HTMT design iteration (which we call COOL-1). In contrast to COOL-0, this design will be based on a special computer architecture [8] and validated by simulation at several design levels. Hopefully, this work will be completed in a year or so, well before the HTMT project moves into its planned Phase 4 (which should include actual building of a computer prototype) in October 2000. We also hope that the COOL-1 design will also become the basis for affordable (personal and corporate) teraflops-scale platforms and servers [2], leading to the eventual commercialization of RSFQ technology.

References

[1] K.K. Likharev, V.K. Semenov, IEEE Trans. Appl. Supercond. 1 (1991) 3.
[2] K.K. Likharev, in: H. Weinstock (Ed.), Applications of Superconductivity, Kluwer, Dordrecht, 1999.
[3] T. Sterling, HTMT computer architecture for petaflops computing, Available on the Web at http://htmt.cacr.caltech.edu/Overview.html, 1998.
[4] M. Dorojevets, P. Bunyk, D. Zinoviev, K.K. Likharev, IEEE Trans. Appl. Supercond. 9 (1999) 3606.
[5] V. Patel, J. Lukens, IEEE Trans. Appl. Supercond. 9 (1999) 3247.
[6] W. Chen, A.V. Rylyakov, V. Patel, J.E. Lukens, K.K. Likharev, IEEE Trans. Appl. Supercond. 9 (1999) 3212.
[7] B. Ruck, R. Schmitz, N. Thyssen, B. Hermanns, H. Kohlstedt, S. Lomatch, IEEE Trans. Appl. Supercond. 9 (1999) 3966.
[8] M. Dorojevets, The COOL-I ISA handbook: Version 1.00, in: Technical Report, RSFQ System Group, SUNY at Stony Brook, January 1999.