Do Superconducting Processors Really Need Cryogenic Memories? The Case for Cold DRAM

Fred Ware, Liji Gopalakrishnan, Eric Linstadt, Sally A. McKee, Thomas Vogelsang,
Kenneth L. Wright, Craig Hampel, Gary Bronner

Rambus

Sunnyvale, CA

{ware, lijig, elinstadt, smckee, tvogelsang, kwright, champel, gbronner}@rambus.com

ABSTRACT

Cryogenic, superconducting digital processors offer the promise of greatly reduced operating power for server-class computing systems. This is due to the exceptionally low energy per operation of Single Flux Quantum circuits built from Josephson junction devices operating at the temperature of 4 Kelvin. Unfortunately, no suitable same-temperature memory technology yet exists to complement these SFQ logic technologies. Possible memory technologies are in the early stages of development but will take years to reach the cost per bit and capacity capabilities of current semiconductor memory. We discuss the pros and cons of four alternative memory architectures that could be coupled to SFQ-based processors. Our feasibility studies indicate that cold memories built from CMOS DRAM and operating at 77K can support superconducting processors at low cost-per-bit, and that they can do so today.

1 Introduction

The steadily shrinking transistor sizes of silicon complementary metal oxide semiconductor (CMOS) technology now let us put over 15 billion transistors on a chip [1]. The energy dissipation of those transistors is reaching physical limits, though. For instance, in 2014 US data centers consumed about 70 billion kilowatt-hours of electricity, and by 2020 usage is projected to reach 73 billion [2]. To address these power problems, scientists and computer designers are turning to "beyond CMOS" technologies such as superconducting single-flux quantum (SFQ) electronics that provide very fast, low-energy switching as well as fast and loss-less interconnects between circuit elements [3].

Hybrid computing systems including quantum components are likely realizable in the near term. Researchers at IBM recently announced that they will build "commercially available universal quantum computing systems" [4] to be deployed through their cloud infrastructure, and the IARPA Cryogenic Computing Complexity (C3) program, aims to demonstrate an energy-efficient, small-scale superconducting computer [5] in the next few years. One of C3's research thrusts is to develop energy-efficient cryogenic memory, and the other is to develop the logic, interconnects, and system plan. The first of these presents the greatest challenges to timely implementation. In fact, the National Security Agency cites the lack of a practical, scalable, 4K memory as one of the main reasons that superconducting digital electronics have only been implemented in niche applications [6].

Proposals for processors built from SFQ circuits uniformly call for memories also operating at 4K [6] [7] [8] [9], largely for the lower power and latency that physical proximity affords. Historically, the path from initial research results to commercial memory products spans many decades, and while resistive/magnetic technologies hold much promise for the future, we believe that future is still many years away. Given this, we consider memory systems built with current semiconductor memory.

To frame our discussion, we first give some background about DRAM and SFQ technologies. We then compare four candidate memory systems either operating at 4K and cooled by liquid helium or operating at 77K and cooled by liquid nitrogen. The first is a future (SFQ/JJ-compatible) 4K memory, the second a similar 4K memory with CMOS support circuitry, the third a 4K CMOS DRAM, and the fourth a 77K CMOS DRAM. With an eye to the future we briefly discuss two other emerging technologies – RRAM and ST-MRAM – and their potential for operation at cryogenic temperatures.

Our analysis indicates that the CMOS DRAM subsystem running at 77K represents a promising design point for realizing superconducting computers in the near term. Modeling, simulation, and experiments with commercial components reveal no significant problems. Cold DRAM memory systems can provide large capacities inexpensively, with reasonable latencies and energy costs.

2 DRAM

We briefly outline DRAM organization. Storage cells are arranged in banks of rectangular arrays connected to peripheral support circuits via word lines and bit lines. The support circuits take an address, decode it, and fan-out the decoded address to many word lines. The storage element selected on each bit line must fan-in with the other, unselected storage elements to a sense amplifier.
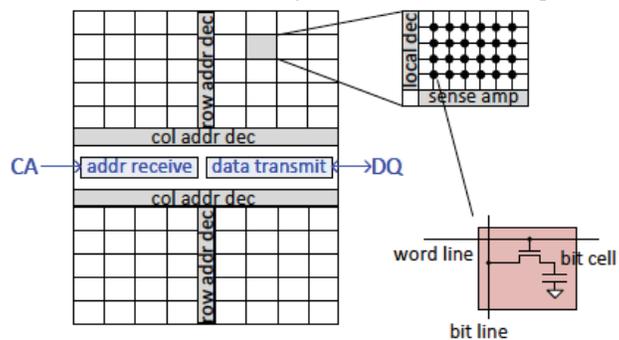

Figure 1: Diagram of DRAM storage array and peripheral circuitry. CA is the clock/command/address signal from the memory controller, and DQ is the bidirectional data bus. Each bit cell contains one transistor and one capacitor (1T1C).

Fig. 1 shows the typical DRAM organization, including 2D arrays of storage cells, sense amps, row and column decoders, and

peripheral circuits. DRAM accesses open a row by loading its entire contents into the sense amps, from which data may be read or written. When different data are needed, the memory controller closes the current row by writing the contents of the sense amps back to the storage array and precharging the bit lines for the next access. Bit charges leak over time, and so the entire DRAM must periodically be refreshed. Refresh simply opens and closes rows without servicing intervening accesses.

Note that DRAM memory systems currently account for up to 40% of total system power and up to 60% of the power for each CPU+DRAM processor node [10] [11], and this is predicted to grow to about 75% of processor-node power budgets in exascale systems [12]. Refresh constitutes a significant portion of memory power for many workloads. Liu et al. project that in four generations, almost half of a room-temperature memory's contribution to system power will come from refresh alone [13]. Operating at colder temperatures dramatically decreases these costs: leakage goes down by roughly an order of magnitude with every 30K decrease in temperature.

3 Single-Flux Quantum Logic

Most proposals for superconducting processors rely on circuits composed of some form of Josephson-junction (JJ) logic [14]. Josephson junctions place a thin, tunneling barrier between two superconducting electrodes – e.g., two Niobium plates separated by an aluminum oxide barrier. At temperatures near absolute zero, pairs of superconducting electrons tunnel through the barrier from one terminal to the other with no resistance, resulting in zero voltage drop. Unlike CMOS circuits implemented with transistors, a JJ acts like a non-linear inductance that accumulates energy when a current is passed through it. Applying a sufficient current to the junction causes it to stop superconducting, and an AC voltage develops across it. In other words, below the critical current threshold, the voltage is zero, and above it, the voltage is non-zero, which makes JJs suitable for building digital circuits.
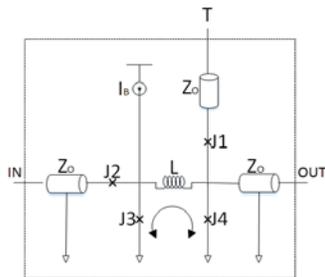

Figure 2: This JJ circuit implements the equivalent of a CMOS flip flop.

Digital circuits that operate near absolute zero can be built with Josephson junctions, Single Flux Quantum voltage pulses, and superconducting transmission lines. The basic energy storage elements are capacitors and inductors. In a JJ circuit, a current pulse I circulating around a superconducting loop with inductance L produces a magnetic flux quantum ϕ according to the equation $\phi=L*I$. When the current exceeds the critical threshold I_c, an SFQ voltage pulse is generated. The quantized pulse is $h/2e$, where h is Planck's constant and e is the electron charge. This pulse is about $2*10^{-15}$ volt-seconds, or 2 millivolt-picoseconds (2mV·ps). The SFQ pulse can propagate in a controlled manner through a series of Josephson

junctions and superconducting loops, or it can be stored within a loop by adjusting the inductance of the loop and injecting an appropriate bias current.

Fig. 2 shows a digital latch built from four JJ devices (J1, J2, J3, and J4), an inductor L, and a pulsed power supply (current source I_B) [15]. An SFQ set pulse injected at the input (IN) is transferred across the Josephson junction J2 (the inductance of that transfer loop is designed to be very low and hence is not shown in the figure). The bias current I_B causes the total current in J3 to exceed the critical current I_c, switching J3 and transferring the pulse into the inner loop. This loop is designed with high inductance L, which causes the current to keep circulating and energy to be stored like a memory cell or flip flop. When a reset SFQ pulse is applied at T, current flows through both J1 and J4 to ground. If the flip flop is in state 0, the smaller junction J1 switches while the voltage across J4 – and hence the output – remains zero. If in state 1 (SFQ stored in L), the larger junction J4 switches, sending the stored SFQ pulse to the output and resetting the latch to the logic 0 state.

Due to the extremely tiny and narrow nature of the SFQ voltage pulses and the superconducting propagation, energy dissipation is negligible compared to conventional CMOS circuits. On the other hand, the tiny pulses and low isolation provided by Josephson junctions make fan-in or fan-out challenging. Amplification is also difficult. JJs are thus unsuitable for driving bit lines and word lines or for implementing the sense amplifiers in memory arrays.

4 Candidate Memory Architectures

We next compare four memory subsystems for a processor utilizing SFQ/JJ digital logic in a 4K temperature domain. The memories are tightly coupled to the processor.
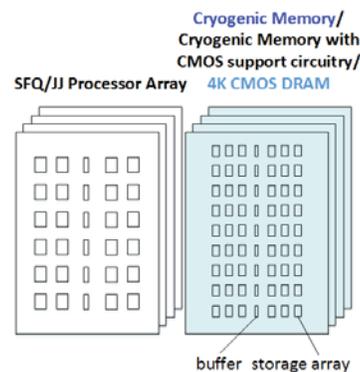

Figure 3: The left side shows the superconducting processor, and the right side represents any of our three 4K memory subsystems.

Fig. 3 shows the high-level organization of a superconducting computing system with a memory subsystem operating within the same 4K temperature domain. The memory would use signaling levels that are compatible with the processor, greatly reducing energy per access versus conventional semiconductor memories.

4.1 Emerging 4K Memory

Several potential technologies are being investigated for memories operating at 4K. The IARPA C3 program is developing orthogonal spin transfer torque cells optimized for operation at 4K [7], cryogenic spin Hall effect cells [16], and cells based on Josephson junctions with properties modified by magnetic layers [17]

[18]. Other hybrid approaches integrate magnetic spin valves with JJs [19]. Hafnium Oxide-based Resistive RAM (RRAM) has recently been shown to operate properly in the 4K domain [20]. (We discuss RRAM further in Section 5.1.) Memories constructed from superconducting nanowires have also been proposed [21]. Unfortunately, all these technologies are still under development: none yet provides the combination of large capacity and low latency required to support a superconducting processor. Furthermore, such memories are unlikely to match the capacity and cost per bit of CMOS DRAM memory for some time.

4.2 4K Memory+CMOS

Commercially available CMOS components have been demonstrated to function correctly when operating at 4K [22]. CMOS circuitry could thus be employed for the support circuitry in a 4K memory array. Such a component would look much like Figure 1 with the 1T1C DRAM cells replaced by bit cells in these new technologies. Operation is similar to that of DRAM, with the CMOS circuits efficiently managing the required fan-out and fan-in. The memory devices need not be directly compatible with the processor's SFQ circuits. A CMOS buffer component performs conversion between SFQ signals and the CMOS circuitry.
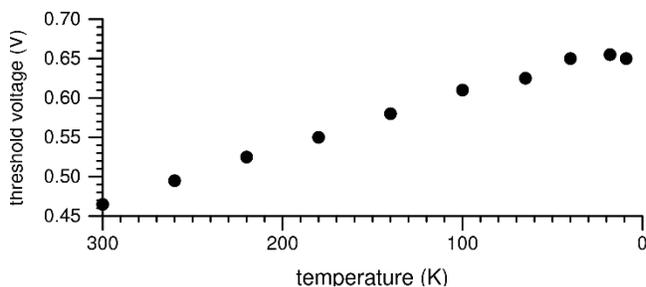

Figure 4: Xie et al. [24] measure threshold voltage for H-gate devices implemented in a 0.18μm PD-SOI process, finding that V_{th} rises almost linearly by about 40% from room temperature down to about 40K, at which point V_{th} becomes nearly constant. (Note the decreasing temperature on the x axis.)

Threshold voltages increase at colder temperatures. Xie et al. [24] measure threshold voltages at decreasing temperatures for H-gate devices implemented in a 0.18um partially depleted SOI process. Fig. 4 shows that they observe a steep gradient until temperatures reach about 40K, but V_{th} is nearly constant below that.

The CMOS will need power supply headroom only slightly less than that required at room temperature, and the power supply will be many orders of magnitude larger than a power supply used by a cryogenic memory device. With the most aggressive voltage possible for the 4K CMOS circuits, we estimate the VDD supply voltage to be about 400mV. This imposes a minimum access energy per bit of about 0.5pJ. Cooling inefficiencies further diminish potential energy benefits of emerging, 4K-temperature memories.

4.3 4K CMOS DRAM

Employing CMOS memory devices modified to operate in the 4K temperature domain is possible in a shorter time frame with fewer technical risks than employing the alternatives discussed above. Lowering operating temperature decreases leakage and refresh power requirements because it exponentially decreases the number of thermally excited charge carriers crossing a semiconductor p-n junction. In addition, switching energy decreases because transistor subthreshold swing gets steeper, allowing on-off control with smaller voltage swings. Reduced leakage current and better transistor switching allows operating voltages to be reduced with lower temperature, but voltage scaling is still limited by temperature-independent mechanisms like coupling between neighboring lines in a dense array.

Bipolar circuits have problems with emitter injection efficiency at cryogenic temperatures, but such circuits (e.g., the bandgap voltage reference) can simply be replaced by others that do not rely on bipolar transistors. Other modifications may also be needed to ensure device reliability and for efficient testing during manufacturing, especially when testing at different temperatures from the targeted operating temperature.

Putting the DRAM in the 4K domain also improves latency because of the close physical proximity to the JJ logic. As with the hybrid system, though, little voltage scaling is possible for a CMOS memory operating at 4K: it, too, requires a supply voltage of about 400mV, causing access energy to be about 1pJ per bit.
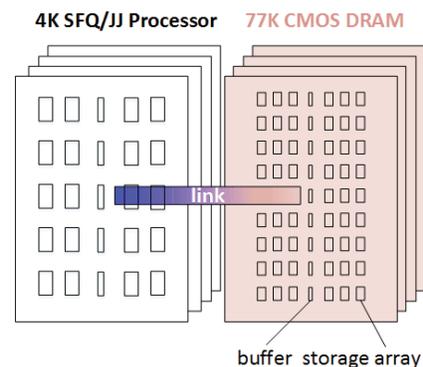

Figure 5: The left side of this diagram again shows the superconducting processor running at 4K, and the right side represents the CMOS memory subsystem running at 77K. The narrow rectangles on both sides represent the buffers necessary to translate between the temperature domains.

4.4 77K CMOS DRAM

Operating CMOS DRAM memories at 4K incurs cooling cost approximately 200 times those at room temperature. Placing CMOS memory devices in an intermediate temperature domain provides the same benefits as in the 4K domain: negligible leakage, large capacity, and low cost. This greatly reduces the impact of cooling inefficiency on the energy per access. Fig. 5 depicts this alternative.

We choose 77K because liquid nitrogen is inexpensive and because the infrastructure for it is widely developed, making this an obvious target domain for optimizing energy expenditure as well as capital and operating expenses. The 77K memory substrate needs a link interface with low thermal conductivity (e.g., a flexible cable [25]) to connect to the 4K processor substrate. The 4K-to-77K link plus receiver operates in the 77K domain. For clarity, we only show a single cable in Fig. 5. In an actual system, each processor substrate would connect to the adjacent memory substrate with multiple cables extending across the gap between the 4K and 77K domains.

The conductive metal wires of the link will contribute most to the thermal conductivity. The insulating layers should not contribute much additional thermal conduction. Thermal conductivity is not an issue for the 4K memories since they can use a single substrate for the processor-to-memory connections.

To minimize the number of link wires, the link must operate at the highest possible signaling rate (somewhere in the 6.4 to 12.8 Gb/s range). The signaling integrity of the link could become an issue at these data rates. There is a tradeoff between the need for low thermal conductivity of the link material and the need for high electrical conductivity. Making the link wire thinner (more resistive) increases signal attenuation, which could impact the design of the receiver in the buffer component.

Note that the physical layer interface design at either end of the link may pose challenges. For instance, the SFQ pulses must be amplified and sampled at a high data rate as they pass through the 73K temperature differential.

Table 1: The main contributors to 77K DRAM energy are memory access and transport to/from the associated buffer component followed by the primary link and associated physical interfaces. Cooling overheads are not included.

Component	Approximate energy (pJ/bit)
memory access and secondary transport	0.8
voltage amplification, timing adjustment, primary link transport	0.4
thermal conduction in links going from 77K to 4K	0.004
thermal conduction in power wires going from 300K to 77K	0.003
thermal radiation from links going from 77K to 4K	0.0002

We use Vogelsang's memory system architecture and power model [26] to explore power-performance tradeoffs. Table 1 shows a breakdown of some of the energy components for an access in this system. Energy contributions from thermal phenomena are dwarfed by row and column access , by transport to/from the associated buffer component, and by the links and physical interfaces.

Our simulations of DRAM component circuits operating at 77K reveal no significant functional issues, and initial experiments with commercially available devices indicate that cold temperatures do not prevent operation. (Details are beyond the scope of this position paper.)

The energy dissipated at each temperature domain adds heat to that domain, and the energy must be removed from the lower-temperature domains and carried to the ambient domain. A cooling system that moves heat from a domain T0 to a warmer domain T1 has a thermodynamic (Carnot) efficiency of T0/(T1-T0). This gets multiplied by a mechanical efficiency factor, ME, and inverted to give an energy/power cooling coefficient. The ME of commercial cooling systems at 77K is known to be about 0.3, and thus the energy

per bit needed for refrigeration power to remove the energy from 77K to the 300K domain is about 10 (i.e., (300-77)/(77*0.3)) times the energy per bit needed for memory access and transport within the 77K domain. The cost of moving energy from the 4K to the 300K domain is an order of magnitude larger.

5 Other Potential Cold Memories

What about other memory technologies operating in an intermediate temperature domain? DRAM process technology is expected to slow and/or stop scaling within the next few years. Most DRAM vendors now have 20nm DRAM process technology in high-volume manufacturing and have roadmaps to scale to the low teens. The economic benefit of scaling is diminishing because the added process complexity of the lower technology nodes reduces the cost benefit of having more die/wafer. The memory industry is responding to this challenge by investigating new memory technologies. We briefly discuss their potential application in cold (intermediate temperature domain) memory systems.

Two contenders are receiving considerable investment in the industry: resistive RAM (RRAM), and spin torque transfer magnetic RAM (STT-MRAM). We consider whether these have advantages over DRAM for low-temperature operation. One of the issues for these emerging technologies is that their write energy per bit is much higher than DRAM, which is a disadvantage for cryogenic operation.

5.1 RRAM

Resistive RAM (RRAM) is considered a promising emerging nonvolatile memory (NVM) due to its fast, low-power operation and its simple manufacturing processes. RRAM is metal-insulator-metal (MIM) sandwiched between two electrodes. Memory '1' and '0' states can be represented by appropriately applying voltages between the two electrodes to create or disrupt a weak current conduction path in the insulator. This insulator switching was first observed in the 1960s [26], but intensive RRAM research has only taken place over the last 15 years. In 2014, Micron demonstrated a 16 Gb RRAM functional chip implemented in a 27nm node with a DDR interface, 180 MB/s writes, and 900 MB/s reads [27] [28]. RRAM cannot compete with 3D NAND in terms of cost per bit, but it can compete with DRAM.

In the cryogenic Hafnium Oxide-based RRAMs reported recently [20] [29] [30], power consumption goes up as temperature decreases because the SET and RESET voltages increase slightly. In addition, the oxygen ion thermal energy decreases, meaning the oxygen needs a higher electrical field to overcome the barrier to form or disrupt the conductive filament.

RRAM retention times at low temperatures could exceed those at room temperature, since the atoms are less thermally activated. Endurance may decrease slightly due to the lower thermal activated energy of oxygen and the need for higher voltages.

5.2 STT-MRAM

STT-MRAM [31] is a leading candidate to displace SRAM for L2 and L3 caches in mobile processors and SOCs, although intrinsic write latencies of 1-10ns will not support its use as an L1 cache. The technology's non-volatility greatly reduces leakage, and the smaller cell size enables larger caches or smaller die.

Information is stored in the relative orientation of two magnetic layers (one with a fixed or reference orientation and the other layer free to switch its orientation) separated by an insulating (tunnel) barrier. The probability of electrons tunneling across the barrier is greater when the fields of the magnetic layers (as well as the electrons' spins) are aligned. The differences in resistance provide a mechanism to read the state of an STT-MRAM cell. To write a bit, a spin polarized current of sufficient density exerts a torque strong enough to switch the orientation of the free magnetic layer. Writes are probabilistic and require ECC. ECC is also required to handle read disturbs and retention errors.

Cryogenic operation is unlikely to reduce power significantly: there is no inherent voltage scaling because STT-MRAM is based upon tunneling junctions. The lower thermal energy means cell currents and areas can be reduced, offering a potential decrease in read and write power and enabling further scaling of the access transistor dimensions, but this reduces the thermal assistance that initiates switching, which increases cycle times and/or write bit error rates.

6 Discussion

Attempts to build quantum logic circuits based on the low-temperature superconductivity of Josephson junctions have been underway since the 1950s, but we have yet to produce a full-fledged system containing a superconductor (co)processor. The lack of memory devices running at the same temperature as the cryogenic superconducting processors is commonly cited as the major obstacle. Emerging memory technologies operating at 4K offer the potential for faster access times (avoiding the latency of moving bits across temperature domains) and average access energies of 10^{-16} to 10^{-17} joules per bit with no static power dissipation, but to date no such technology has met the combined requirements for density, low power, and high performance.

Historically, new memory technologies have taken many decades to go from the research lab to the marketplace. For instance, PCM has been studied since the 1960s, but the first products did not appear until 2010, and these were not commercially successful. Given the amount of research currently devoted to their development, 4K memories are likely to be available relatively faster, but they are still many years away.

In contrast, cryogenic processors based on SFQ logic will likely be available in the near term. The question, then, is what kinds of memory systems to pair with them. While waiting for the new

memory technologies to mature, we consider the alternative of using DRAM operating at colder temperatures.

Table 2 summarizes tradeoffs among the four alternative memory systems we discuss here. The rightmost column gives approximate energy costs, including cooling. Although employing CMOS support circuits with a 4K cryogenic memory technology removes the requirement that the memory devices be directly compatible with the processor's SFQ circuits, it does so at a high price: the energy costs of the emerging memory with CMOS support circuits are dominated by the CMOS, giving this design point few advantages over a CMOS DRAM system running at 4K. By locating the memory close to the processor, the 4K DRAM system provides low access latency, but it incurs high cooling costs. Placing the DRAM in a 77K temperature domain delivers most of the benefits of the 4K system (negligible leakage, large capacity, and low cost) with much lower cooling costs. Even though DRAMs consume more power than what is expected for the new memory technologies, we can still deliver highly power-efficient systems in the near term because the switching energy for the gates in superconducting processors is an order of magnitude less than CMOS.

References

[1] M. Harris, "Inside Pascal: NVIDIA's Newest Computing Platform," 19 June 2016. [Online]. Available: https://devblogs.nvidia.com/parallelforall/inside-pascal/. [Accessed 20 April 2017].

[2] Y. Sverdlik, "Here's how much energy all US data centers consume," 27 June 2016. [Online]. Available: http://www.datacenterknowledge.com/achives/2016/06/27/heres-how-much-energy-all-us-data-centers-consume/. [Accessed 14 April 2017].

[3] MIT Lincoln Laboratory, "Forecasting Superconductive Electronics Technology," The Next Wave, vol. 23, no. 3, p. 11, 2014.

[4] IBM, [Online]. Available: https://www-03.ibm.com/press/us/en/pressrelease/51740.wss.

[5] M. A. Manheimer, "Cryogenic Computing Complexity Program: Phase 1 Introduction," IEEE Transactions on Applied Superconductivity, vol. 25, no. 3, p. 1301704, March 2015.

[6] F. Bedard, N. K. Welker, G. R. Cotter, M. A. Escavage, and J. T. Pinkston, "Superconducting Technology Assessment," National Security Agency Office of Corporate Assessments, 2005.

[7] L. Ye, D. B. Gopman, L. Rehm, D. Backes, G. Wolf, T. Ohki, A. F. Kirichenko, I. V. Vernik, O. A. Mukhanov, and A. D. Kent, "Spin-transfer switching of orthogonal spin-valve devices at cryogenic temperatures," IOP Journal of Applied Physics, vol. 115, no. 17, p. 17C725, February 2014.

[8] D. S. Holmes, A. L. Ripple, and M. A. Manheimer, "Energy-Efficient Superconducting Computing—Power Budgets and Requirements," IEEE

Table 2: Summary of system tradeoffs. DRAM operating in a 77K temperature domain enables us to build scalable memories for SFQ processors at reasonable latencies and power budgets. DRAM in the 4K domain offers latency advantages, but the power costs are unacceptable. Emerging cryogenic memories will eventually let us build extremely low-power, high-speed memory subsystems, but those technologies are still several years out.

Architecture	Available?	Can Scale to Required Capacity?	Transport Latency (processor cycles)	Approx. energy per access (pJ/bit)
new memory @4K	in research	not yet	~1	0.2
new memory+CMOS @4K	in research	not yet	~2	120
DRAM @4K	soon	yes	~10	240
DRAM @77K	now	yes	~100	12

Transactions on Applied Superconductivity, vol. 23, no. 3, p. 1701610, June 2013.

[9] T. V. Duzer, L. Zheng, S. R. Whitely, H. Kim, J. Kim, X. Meng, and T. Ortlepp, "64-kb Hybrid Josephson-CMOS 4 Kelvin RAM With 400 ps Access Time and 12 mW Read Power," IEEE Transactions on Applied Superconductivity, vol. 23, no. 3, p. 1700504, June 2013.

[10] C. Lefurgy, K. Rajamani, F. Rawson, M. Kistler, and T. W. Keller, "Energy management for commercial servers," IEEE Computer, vol. 36, no. 12, pp. 39-48, December 2003.

[11] L. Barroso and U. Hőzle, "The datacenter as a computer: An introduction to the design of warehouse-scale machines," Synthesis Lectures on Computer Architecture, vol. 4, no. 1, pp. 1-108, 2009.

[12] S. Borkar, The Exascale Challenge, keynote, in International Symposium on VLSI Design, Automation and Test, pp. 2-3, April 2011.

[13] J. Liu, B. Jaiyen, R. Veras, and O. Mutlu, "RAIDR: Retention-aware intelligent DRAM refresh," in International Symposium on Computer Architecture , pp. 1-12, June 2012.

[14] R. Newrock, "What are Josephson junctions? How do they work?" 2017. [Online]. Available: https://www.scientificamerican.com/article/what-are-josephson-juncti/. [Accessed 4 May 2017].

[15] D. S. Holmes, A. M. Kadin, and M. W. Johnson, "Superconducting Computing in Large-Scale Hybrid Systems," IEEE Computer, vol. 48, no. 12, pp. 34-42, December 2015.

[16] L. Liu, C.-F. Pai, H. W. Tseng, D. C. Ralph, and R. A. Buhrman, "Spin-Torque Switching with the Giant Spin Hall Effect of Tantalum," Science, vol. 336, no. 6081, pp. 555-558, 4 May 2012.

[17] B. M. Niedzielski, S. G. Diesch, E. C. Gingrich, Y. Wang, R. Loloee, W. P. Pratt, and N. O. Birge, "Use of Pd–Fe and Ni–Fe–Nb as Soft Magnetic Layers in Ferromagnetic Josephson Junctions for Nonvolatile Cryogenic Memory," IEEE Transactions on Applied Superconductivity, vol. 24, no. 4, p. 1800307, August 2014.

[18] M. A. El Qader, R. K. Singh, S. N. Galvin, L. Yu, J. M. Rowell, and N. Newman, "Switching at small magnetic fields in Josephson junctions fabricated with ferromagnetic barrier layers," Applied Physics Letters, vol. 104, p. 022602, January 2014.

[19] B. Baek, W. H. Rippard, S. P. Benz, S. E. Russek, and P. D. Dresselhaus, "Hybrid superconducting-magnetic memory device using competing order parameters," Nature Communications, vol. 5, pp. 1-6, May 2014.

[20] R. Fang, W. Chen, L. Gao, W. Yu, and S. Yu, "Low-Temperature Characteristics of HfOx-Based Resistive Random Access Memory," IEEE IEEE Electron Device Letters, vol. 36, no. 6, pp. 567-569, June 2015.

[21] S. Pagano, N. Martucciello, F. Bobba, G. Carapella, C. Attanasio, C. Cirillo, R. Cristiano, M. Lisitskiy, M. Ejrnaes, G. P. Pepe, and L. Parlato, "Proposal for a Nanoscale Superconductive Memory," IEEE Transactions on Applied Superconductivity, vol. 27, no. 4, p. 1801004, June 2017.

[22] I. D. Conway, J. I. Colless, J. M. Hornibrook, S. J. Pauka, S. J. Waddy, M. K. Frechtling, and D. J. Reilly, "An FPGA-based instrumentation platform for use at deep cryogenic temperatures," AIP Review of Scientific Instruments, vol. 87, no. 1, January 2016.

[23] T. Chen, C. Zhu, L. Najafizadeh, B. Jun, A. Ahmed, R. Diestelhorst, G. Espinel, and J. D. Cressler, "CMOS reliability issues for emerging cryogenic Lunar electronics applications," Solid-State Electronics, vol. 50, no. 6, pp. 959-963, June 2006.

[24] B. Xie, B. Li, J. Bi, B. Bu, C. Wu, B. Li, Z. Han, and J. Luo, "Effect of cryogenic temperature characteristics on 0.18μm silicononinsulator devices," Chinese Physics B, vol. 25, no. 7, p. 078501, 2016.

[25] T. Vogelsang, "Understanding the Energy Consumption of Dynamic Random Access Memories," in International Symposium on Microarchitecture, pp. 363-374, December 2010.

[26] J. F. Gibbons and W. E. Beadle, "Switching properties of thin NiO films," Solid-State Electronics, vol. 7, no. 11, pp. 785-790, November 1964.

[27] R. Neale, "Showtime for the Micron-Sony 16Gb ReRAM," EE Times, 4 March 2014. [Online]. Available: http://www.eetimes.com/author.asp?doc_id=1321276. [Accessed 4 May 2017].

[28] S. Sills, S. Yasuda, A. Calderoni, C. Cardon, J. Strand, K. Aratani, and N. Ramaswamy, " Challenges for high-density 16Gb ReRAM with 27nm technology," in IEEE Symposium on VLSI Circuits, pp. T106-T107, June 2015.

[29] S. Blonkowski and T. Cabout, "Bipolar resistive switching from liquid helium to room temperature," IOP Journal of Physics D: Applied Physics, vol. 48, no. 34, p. 345101, September 2015.

[30] C. Walczyk, D. Walczyk, T. Schroeder, T. Bertraud, M. Sowinska, M. Lukosius, M. Frashke, D. Wolansky, B. Tillack, E. Miranda, and C. Wenger, "Impact of Temperature on the Resistive Switching Behavior of Embedded HfO2-Based RRAM Devices," IEEE Transactions on Electron Devices, vol. 58, no. 9, pp. 3124-3131, September 2011.

[31] K. L. Wang, J. G. Alzate, and P. K. Amiri, "Low-power non-volatile spintronic memory: STT-RAM and beyond," IOP Journal of Physics D: Applied Physics, vol. 46, no. 7, pp. 074003:1-10, January 2013.