

IARPA Digital Superconducting Circuit Fabrication (C3)
(Lincoln Program #2700)
Final Report

Leonard Johnson, Program Manager

Quantum Information & Integrated Nanosystems Group
MIT Lincoln Laboratory

September 16, 2020

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

This research was funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA) under Air Force Contract No. FA8702-15-D-0001. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of ODNI, IARPA, or the US Government.

© 2020 Massachusetts Institute of Technology.

Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.

I. Background

Today's state-of-the-art computer systems are a result of steady predictable scaling of silicon CMOS (complementary metal-oxide semiconductor) integrated circuit technology, as shrinking transistor dimensions over the past several decades have enabled a more than 10,000x increase in the number of transistors on a single processor chip. However, the energy dissipation of CMOS transistors is reaching physical limits, and has become a difficult barrier to building more powerful supercomputers. Advances in "beyond CMOS" device technologies are now seen as a key step towards achieving the next major leap in high-performance computing.

One candidate for beyond-CMOS technology is digital integrated circuits based on superconductive single-flux-quantum (SFQ) logic. This technology offers a combination of high speed and ultra-low power dissipation unmatched by any other device. First pioneered at Moscow State University in the 1980's by a team led by Professor Konstantin Likharev (the team relocated to Stony Brook University in the early 1990's), SFQ technology has seen a resurgence to address the needs of exascale computing.

Superconductive integrated circuits rely on the fact that magnetic flux is quantized in superconducting wiring loops. In active circuits, these flux quanta are manipulated to perform logic operations by using Josephson junctions (JJs), two superconductors separated by a thin insulating barrier. When the current through a JJ exceeds its critical current, I_C , a single flux quantum (SFQ) pulse is generated. This is an exceptionally fast, low-energy switching process—about 1 picosecond (ps) and 10^{-19} joules (J).

Under the IARPA-funded Cryogenic Computing Complexity (C3) program MIT Lincoln Laboratory (MIT-LL) developed state-of-the-art integrated circuit fabrication capability for the development and demonstration of SFQ-based integrated circuits for high-performance computing. With this microfabrication process capability, MIT-LL performed research foundry fabrication of circuits designed by C3 design team performers. In addition, MIT-LL developed, and provided as a research foundry capability for a superconducting multi-chip module (S-MCM) process comprised of passive superconducting wiring layers on a silicon substrate. Using flip-chip attachment, this process provides lossless interconnection between individual SFQ chips.

The C3 program kickoff meeting was held by IARPA on 29 July 2014, and the final C3 performers meeting was held by IARPA on 3 December 2018.

II. SFQ5ee fabrication process

The vast majority of the C3-funded research foundry SFQ fabrication activity was centered on the SFQ5ee technology node. The MIT-LL fabrication technology is centered on Nb-based Josephson junctions (JJ's) and passive Nb wiring layers, all fabricated on 200-mm silicon substrates. Key process parameters for SFQ technology nodes include the Josephson junction (JJ) critical current density J_C , the number of passive Nb wiring layers, the minimum wiring linewidth, minimum JJ size, and the shunt resistor sheet resistance R_s . The key parameters of the SFQ5ee node, and the prior SFQ3ee and SFQ4ee nodes utilized at the start of the C3 program, are summarized in Table 1.

SFQ Fabrication Technology Nodes

Fabrication Process Attribute		Units	Process Node		
			SFQ3ee	SFQ4ee	SFQ5ee
Critical current density		$\mu\text{A}/\mu\text{m}^2$	100	100	100
JJ diameter (surround)		nm	700	700	700
Nb metal layers		-	4	8	8
Line width (space)	Critical layers	nm	500 (1000)	500 (700)	350 (500)
	Other layers	nm			500 (700)
Metal thickness		nm	200	200	200
Dielectric thickness		nm	200	200	200
Resistor width (space)		nm	1000 (2000)	500 (700)	500 (500)
Shunt resistor value		Ω/sq	2	2	2 or 6
m Ω resistor		m Ω	-	-	3 - 10
High kinetic inductance layer		pH/sq	-	-	8
Via diameter (surround)		nm	700 (500)	700 (500)	600 (300)
Via type, stacking		-	Etched, Staggered	Etched, Stacked I2I	Etched, Stacked I2I

Changes from the previous process

Table 1. Key parameters for SFQ integrated circuit fabrication technology nodes at MIT-LL.

The schematic cross-section of the SFQ5ee node is shown in Fig. 1. The JJ layer is located close to the top of the stack, with most of the passive Nb wiring layers below the JJ. A high-kinetic-inductance layer (HKIL), available as an option, is located at the bottom of the stack.

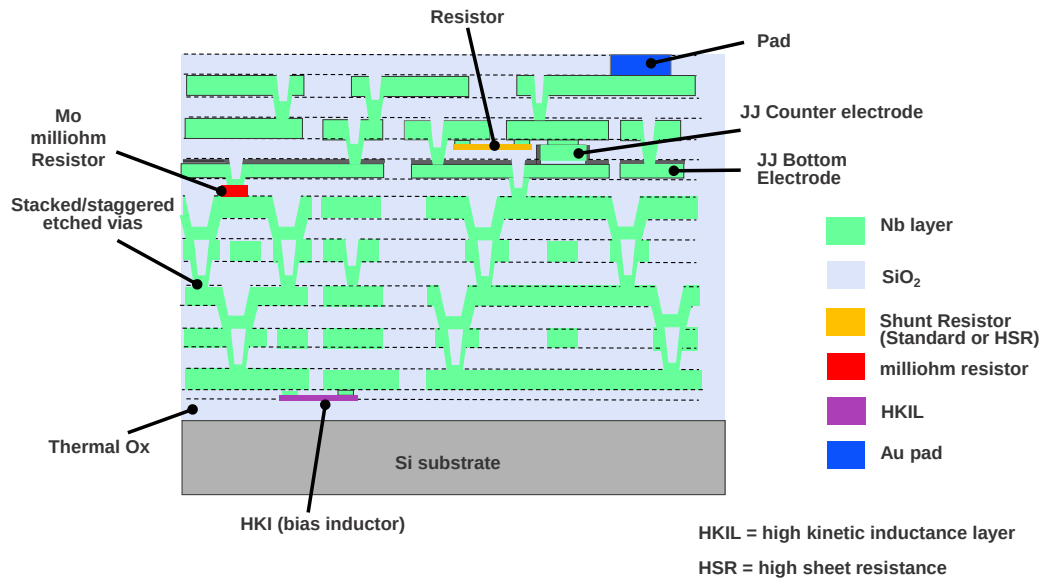


Fig. 1. Schematic cross-section of the SFQ5ee technology node.

Under the C3 program, fabrication was performed in the MIT-LL Microelectronics Laboratory (ML) (Fig. 2). This Class-10 cleanroom facility has a production-class toolset for 200-mm wafer processing. The facility is shared across a wide range of microfabrication projects including Si CMOS, CCD imagers, and integrated photonics. The ML operates on a 5-day/24-hour schedule, initiated under the C3 program. During the C3 program 248-nm photolithography was utilized for most process steps with typical minimum feature sizes of 350 nm. Sputter deposition was employed for niobium (Nb) JJ and wiring layers. The SFQ5ee node was based on etched vias. Chemical mechanical planarization (CMP) is employed throughout the process to maintain uniform dielectric thickness. A microphotograph of a cross-section of the SFQ5ee node is shown in Fig. 3.

MIT-LL Microelectronics Laboratory



Fig. 2. Research foundry SFQ fabrication was carried out in the MIT-LL Microelectronics Laboratory.

SFQ-5ee (9-Metal-Layers)

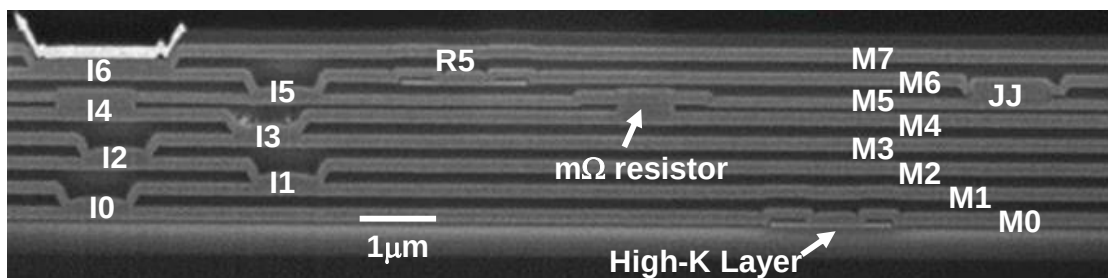


Fig. 3. Microphotograph of a cross-section of the SFQ5ee process node.

III. MIT-LL SFQ5ee research foundry description

Under the C3 program, MIT-LL provided research foundry capability in the SFQ5ee process to the C3 design teams on a multi-project basis. A typical SFQ5ee fabrication run included eight 200-mm wafers. Two photomask (reticle) sets were utilized on each wafer, as shown in Fig. 4. For each mask set, the reticle size was approximately $20 \times 20 \text{ mm}^2$, with each reticle divided into sixteen $5 \times 5 \text{ mm}^2$ chips. The performer reticle included 15 chip designs submitted by performers, and one chip for MIT-LL process monitoring purposes. In addition, there was a separate process control monitor (PCM) reticle set, designed by MIT-LL, which was repeated 9 times across the wafer to provide process statistics.

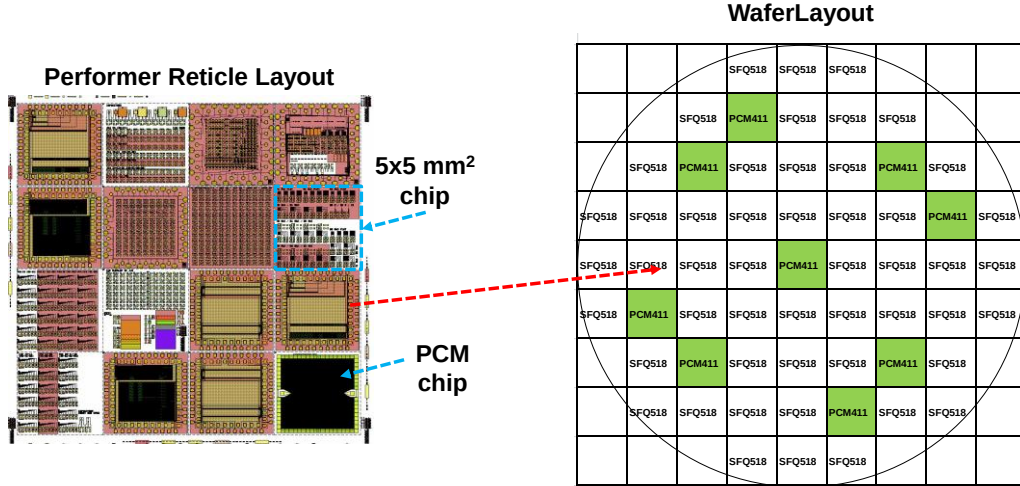


Fig. 4. Mask and reticle layout for C3 SFQ5ee multi-project fabrication runs.

The MIT-LL PCM structures are designed for both room-temperature testing at the wafer level, utilizing automated wafer probers, and for cold (4K) chip testing in a liquid He dewar. Key information on 4K operation can be inferred from room-temperature resistance measurements, and thus is an important tool for obtaining process statistics. A description of the structures included for 4K testing on the PCM reticle set is shown in Fig. 5. Test structures include Josephson junctions, shunt resistors, inductance characterization circuits (SQUID's), and passive wiring and via test structures.

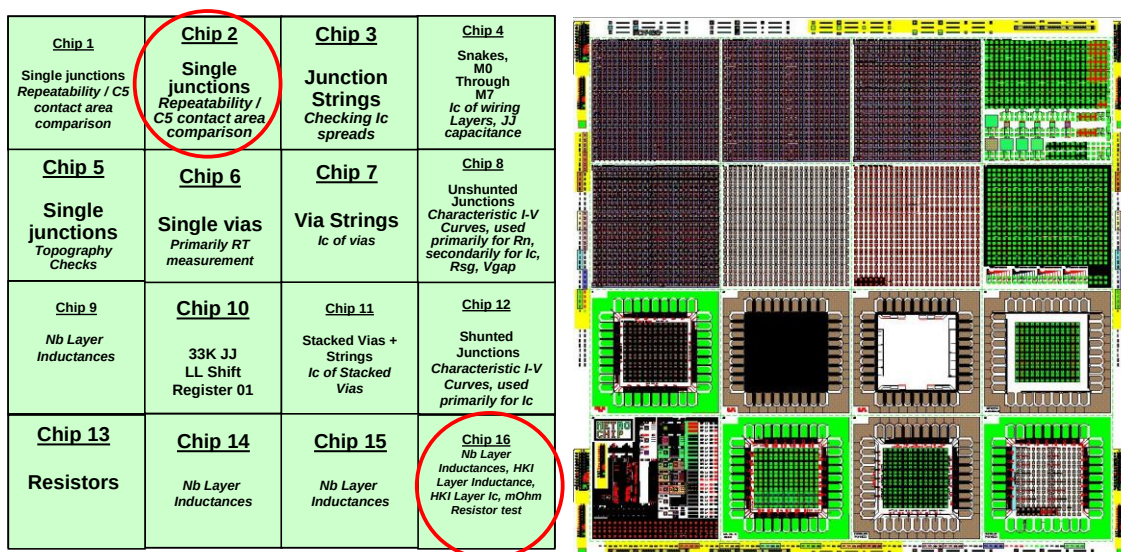


Fig. 5. Description of PCM reticle test structures for cold (4K) process monitor testing.

The timeline for a typical SFQ5ee research foundry fabrication run is shown in Fig. 6. Outside teams design circuits in accordance with design rules provided by MIT-LL. For C3, design submission dates were set by IARPA. Over the course of the program design submissions took place on an approximate quarterly (every three months) basis. Following design submission from various teams, MIT-LL layout personnel performed fill as necessary, and then aggregated designs for the multi-performer reticle set. Actual SFQ5ee wafer fabrication was completed in about 11 weeks. Following wafer fabrication, and before dicing, room-temperature PCM testing was carried out on automated wafer probe stations, and test data was analyzed. The best wafers were then selected for dicing. After dicing, selected 4K PCM testing was done at the chip level. After a final review of PCM test results, chips were delivered (see Fig. 7) to design teams for C3 circuit testing.

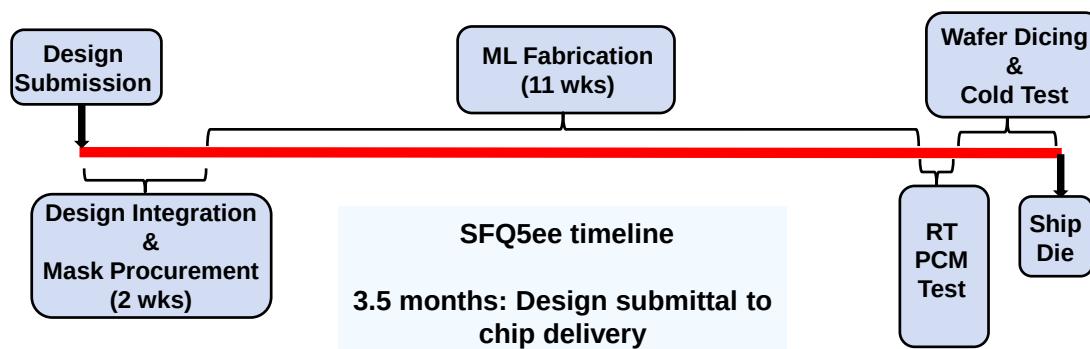


Fig. 6. Timeline for typical SFQ5ee research foundry run, from design submission through initial chip (die) delivery.



Fig. 7. Diced chips ready for delivery to C3 design teams.

IV. Room-temperature PCM testing

As described above, all SFQ5ee fabrication runs performed under C3 included a broad assortment of PCM test structures, repeated across each wafer. These PCM test structures were utilized extensively for process monitoring, parameter targeting, and process troubleshooting when necessary. One of the key test structures were arrays of JJ's designed specifically for room-temperature testing. The measured room-temperature resistance of a JJ, combined with in-line critical dimension (CD) measurements, provides a means to infer 4K JJ parameters including critical current I_C , and critical current density J_C . In addition, by measuring large numbers of nominally identical JJs, the spread in I_C values can be inferred, an important parameter for circuit design. Also, PCM data can be readily used to characterize cross-wafer parameter targeting.

An example of PCM RT testing of arrays of 700-nm JJs is shown in Fig. 8. As explained above, the PCM reticle is repeated 9 times across each SFQ5ee wafer. Each PCM reticle contains a chip (Chip 2) containing, along with other structures, 110 nominally identical 700-nm JJs. Histograms of room-temperature resistance is shown for each die location across the wafer. At the chip level the spread in RT resistance is $\sim 2\%$. Combined with 4K data, the room-temperature measurements can also be used to create a J_C wafermap. An example of such a map is shown in Fig. 9, showing, in this case, that J_C varies within the range of approximately 87 to 95 $\mu\text{A}/\mu\text{m}^2$. (within the target range for a nominal 100 $\mu\text{A}/\mu\text{m}^2$ wafer).

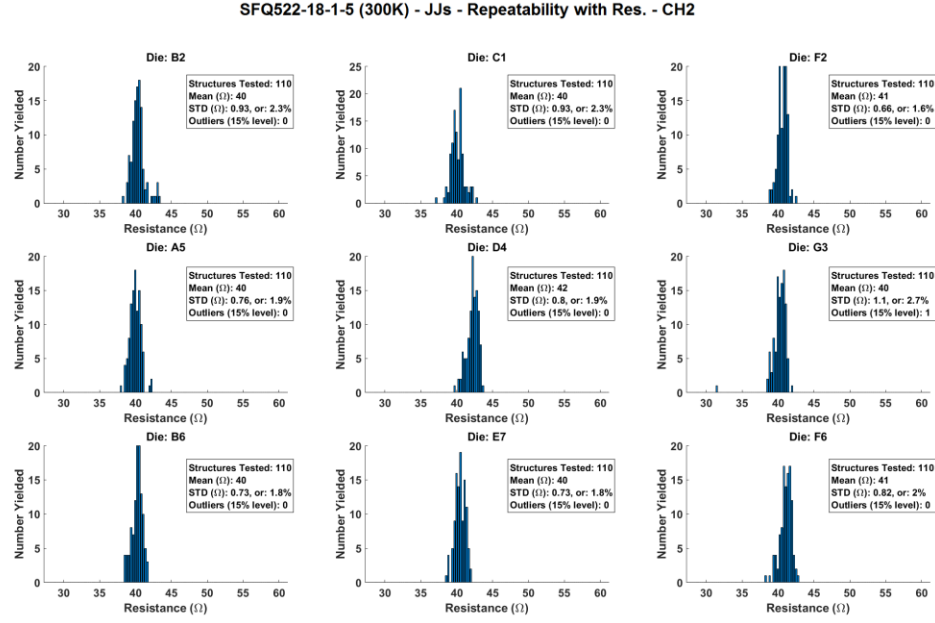


Fig. 8. Example of histograms of measured room-temperature resistance of 700-nm JJs PCM structures. There are 110 JJ's measured for each of 9 wafer locations. The measured spread of resistance values at the chip level is $\sim 2\%$.

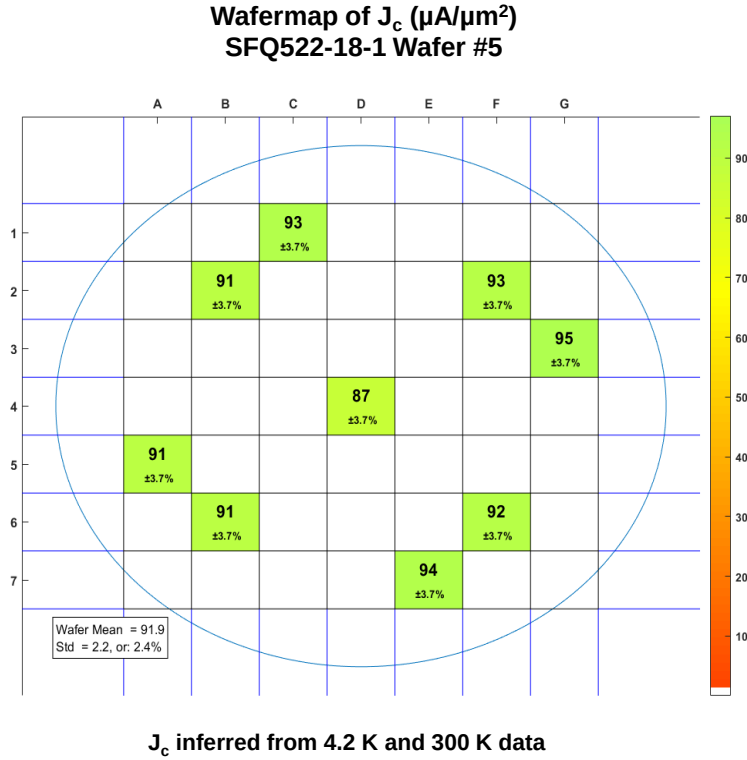


Fig. 9. Example wafer map of critical current density J_c inferred from room-temperature and 4K PCM data.

V. Defect monitoring and defect-free chip yield

A key objective of the C3 program was the demonstration of large-scale (10's K JJs) circuits, which can occupy chip areas as large as $\sim 1 \text{ cm}^2$. From a fabrication perspective, this required close monitoring and reduction of defect levels. As part of the C3 effort, MIT-LL procured and installed a KLA Altair 8920 defect inspection and analysis tool (Fig. 10) specifically for SFQ research foundry purposes. This tool has both brightfield and darkfield capability, and is able to reliably detect defects down to the $\sim 200\text{-nm}$ size level. Once this tool was installed and qualified, it was utilized on all SFQ5ee research foundry fabrication runs, and at all key process levels. Under the C3 program, the overall defect level (killer defects) was brought down to approximately 0.7 killer defects per cm^2 . Reducing particle contamination during film deposition and CMP was essential for reaching this defect level. At the performer chip level, defect-free yields of $>90\%$ were routinely achieved. An example of the wafer-by-wafer chip yield for an SFQ5ee research foundry run (SFQ524) is shown in Fig. 11, indicating defect-free chip yields above 93% for all wafers.

MIT-LL also employs shift registers as a means to characterize yields of working chips based on electrical measurements. As part of this effort under C3, a shift register with 810K JJs on a $1 \text{ cm} \times 1 \text{ cm}$ chip was successfully demonstrated. This, and subsequent versions, represent the world record in JJ count for SFQ circuits, and the highest JJ density of 1.3×10^6 JJs per cm^2 .

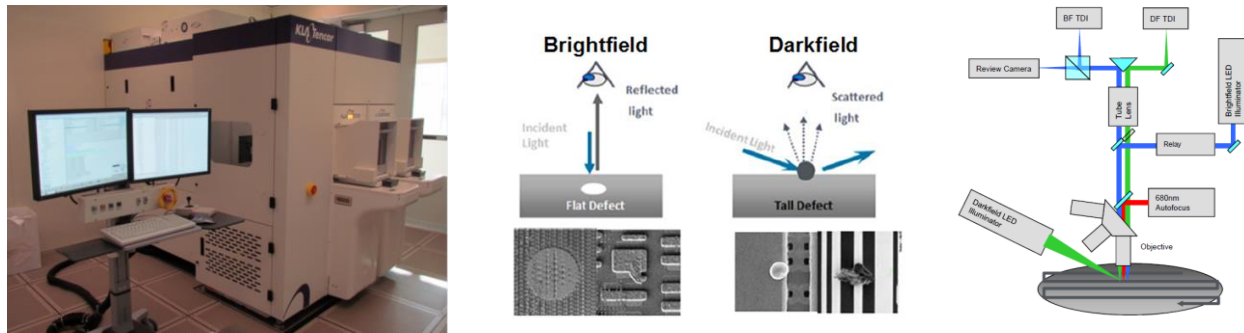


Fig. 10. KLA Altair 8920 defect inspection and characterization tool.

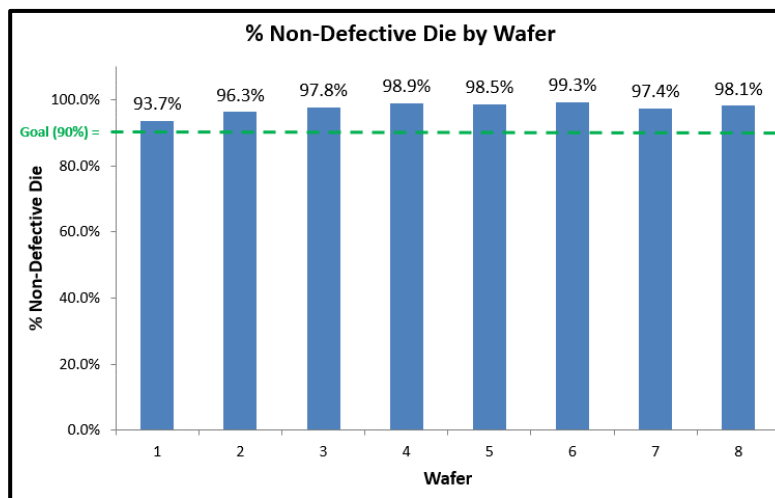
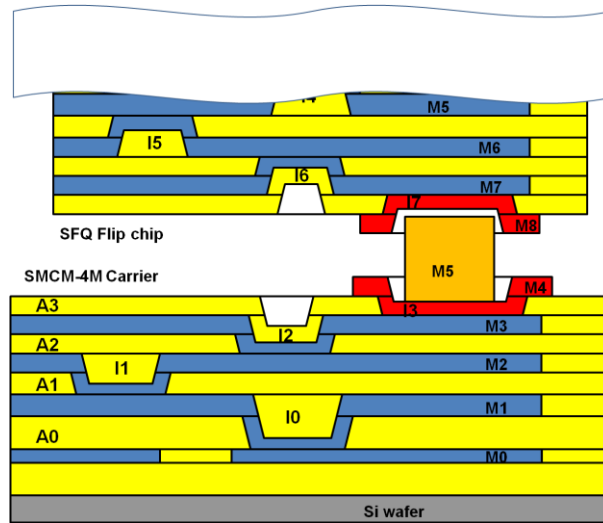


Fig. 11. Wafer-by-wafer defect-free chip yield on fabrication run SFQ524-18-1, completed in 2018.

VI. S-MCM flip-chip packaging

For SFQ system applications it is highly advantageous to interconnect multiple SFQ chips with superconducting wiring in a multi-chip module format. For this purpose, MIT-LL has developed an S-MCM fabrication process which provides (see Fig. 12) four Nb wiring layers on a silicon substrate. The S-MCM process utilizes i-line photolithography and can thus support S-MCM's up to as large as $40 \times 40 \text{ mm}^2$, limited by the i-line reticle size. Multiple SFQ chips can be attached to the S-MCM using a flip-chip attachment process, implemented with In bumps with pitch as low as $35 \text{ }\mu\text{m}$ (see Fig. 13). An example of an S-MCM assembly completed under C3 is shown in Fig. 13, where nine $5 \times 5 \text{ mm}^2$ SFQ chips are interconnected on an S-MCM substrate. Chip-to-chip SFQ communication was demonstrated with this assembly by a C3 design teams.



VII. Process development for future nodes

Under C3, MIT-LL performed exploratory work on developing process modules for future technology nodes. Reducing feature sizes, while maintaining necessary parameter targeting and sufficiently tight parameter spreads, is essential for increasing the density (JJ's per cm^2) of SFQ circuits. One example, as shown in Fig. 14, is the development of a 250-nm-linewidth Nb wiring process, with 200-nm Nb layer thickness. Numerous test structures were evaluated, and results show linewidth variation of $\sim 4\%$, with critical currents I_C of ~ 25 mA, both meeting requirements, thus indicating that a 250-nm node is nearly ready for research foundry availability. Early work, also with promising results, was carried out on self-shunted JJ's, which eliminate the need for external shunt resistors which can otherwise place a limit on achievable circuit densities.

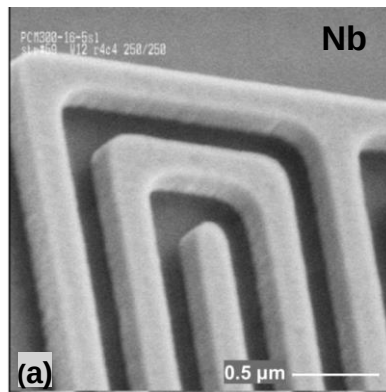


Fig. 14. Photomicrograph of 250-nm-wide 200-nm-thick Nb wiring test structure.

IX. Summary

Under the IARPA C3 program, MIT-LL provided research foundry capability for both SFQ superconducting integrated circuits, and S-MCM SFQ chip packaging. This capability enabled significant SFQ circuit demonstrations by C3 performer (design) teams. Exploratory work on new fabrication process modules, aimed at further increasing SFQ integration scale, has been completed to help guide the specification of future SFQ process nodes.