Single Flux Quantum (SFQ) Circuit Fabrication and Design: Status and Outlook


ORNL Beyond CMOS Workshop

6 April 2016
SFQ Fab is **Not** Dumbed-Down CMOS Fab

- **Looks easy! 1990’s CMOS backend**
  - Feature sizes > 0.25 µm
  - 8 metal layers
  - Etched vias
  - No CMOS front end

- **Only needs 1990’s tool set**
  - Sputtering tool
  - Dielectric deposition
  - CMP
  - Plasma etch tool
  - 248-nm lithography tool

- **But it is not**
  - Circuit performance critically dependent on component values
  - Metal & dielectric thickness to < 10%
  - Junction uniformity < 1% (1σ)
  - Process temperatures < 150 °C
  - Hydrogen incorporation is pernicious
SFQ Circuit Technology is Immature

Superconducting electronics will not succeed if we view Josephson junctions as a drop-in replacement for CMOS transistors

- Cold, hard facts
  - As posed today, SFQ will not have the same density as CMOS circuits
  - Have not developed an optimal circuit design approach
  - Need an effective solution for data storage

- Cold, hard facts (Part 2)
  - SFQ information process is uniquely energy efficient
  - SFQ circuits can be designed for 10X higher clocks than CMOS
  - SFQ circuits have access to a quantized reference (fluxon)
    - Unique advantage for mixed-mode signal processing

Need to develop a computing paradigm that takes advantage of SFQ strengths
Outline

• Short history of SFQ fabrication and circuit develop
• State-of-the-art SFQ fabrication
• Scaling limitations and EDA needs
• Summary
1\textsuperscript{st} Gen SFQ: Latching Logic
IBM 1970’s to early 1980’s

• Logic gate formed from interferometric circuits containing several Josephson junctions
  – Typically “AND” and “OR” gates
  – Designed to latch a voltage state
  – Bipolar power supply served as clock and data reset

• Fabrication evolved to Nb base electrode with Pb-In-Au counter electrode
  – Junction formed on the edge of the Nb electrode
  – Minimum feature size was 2.5 \( \mu \text{m} \)

• Cross Sectional Model (CSM) experiment
  – 10 logic levels, chip-to-chip data flow
  – 300 MHz clock

• Project halted in 1983
  – No path to clocks > 1 GHz
  – No viable memory
  – Continued advance of Si circuits
Beyond CMOS - 6

2nd Gen: Rapid Single Flux Quantum (RSFQ) 1990 – 2010’s

- RSFQ design approach invented in 1985 at Moscow State University
  - Several junctions per logic gate
  - Bits encoded in short voltage pulses
  - DC current bias

- Fabrication based on Nb-Al-AlOx-Nb trilayer junctions
  - Invented in 1971 at IBM
  - 4 Nb-layer standard process
  - Additional layers added in last decade

- Many significant circuit demonstrations
  - Single gate clock speed, 750 GHz
  - 8-32 bit adders, multipliers, ALU at clock speeds of 20 – 30 GHz
  - Analog-to-digital converters with sample rates >40 GHz

- Circuits limited to ~20k JJs

Hypres 4-Metal Layer Process

- 4 Nb wiring layers
- Feature size > 1μm
- Typical junction $I_c$ 100μA, $J_c$ 1 – 4.5 kA/cm²

8-bit ALU with 20 GHz Clock

(Filippov, 2012)
3rd Gen: No Bias Resistor Approaches
RQL, eRSFQ, eSFQ

Conventional RSFQ Biasing

Resistive biasing dominates power dissipation

- DC bias with minimal dissipation: in JJ shunt resistors only during switching
- Can adapt standard RSFQ gates to eSFQ

eSFQ Circuit Design Approach (Hypres)

- Replace R in bias circuit with L
- 0.8 aJ/bit eSFQ shift register and deserializer test circuits*

Reciprocal Quantum Logic, RQL (Northrup Grumman)

- AC Clock, no dc bias resistor dissipation
- Inductively coupled to RF signal line to power devices / eliminates ground return
- Four-phase clock to provide directionality of the SFQ pulse propagation
- “1”s are encoded as a reciprocal pair of SFQ pulses

Schematic of RQL Shift Register Bit

- ~0.1 aJ/bit 8-bit Kogge-Stone RQL adder
- 6.2 GHz clock

*Hypres 4.5 kA/cm²
4th Gen: Adiabatic Quantum Flux Parametron (AQFP)

- The bit is created with two connected loops, each with 1 JJ, inductively coupled to the control lines
  - Excitation current creates double-well potential
  - States ‘0’ and ‘1’ correspond to an SFQ stored on either side of circuit / double well potential

- In adiabatic operation, manipulate information without switching the JJs
  - ‘Gradual’ rise/fall of excitation current for adiabatic operation
  - Switching energy $< I_c\Phi_o$, can be at Landauer limit
  - In contrast, 3rd gen SFQ switches JJs – energy per bit $\sim I_c\Phi_o = 1\text{aJ}$

- Several recent demonstrations confirming the viability of the circuit design approach
  - Logic cells
  - Adder circuits with up to 20k JJs

8-Bit Kogge-Stone Adder

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Kogge-Stone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fab Process</td>
<td>STP2 2.5 kA/cm²</td>
</tr>
<tr>
<td>Number JJs</td>
<td>1224</td>
</tr>
<tr>
<td>Circuit area</td>
<td>1.74 x 0.99 mm²</td>
</tr>
<tr>
<td>Clock</td>
<td>5 GHz</td>
</tr>
<tr>
<td>Latency</td>
<td>1400 ps</td>
</tr>
</tbody>
</table>
| Energy dissipation | 16.4 aJ for $\beta_c = 1.0$
|                    | 10.9 aJ for $\beta_c = 5.0$ |
Strategy for Advancing SFQ Fabrication

Increasing clock speed

Increasing metal layer count and decreasing feature size

Higher $J_C$ (self-shunting junctions)

4 metal layers, Decreased feature size

Estimated Circuit Density (JJ Count)

Critical Current Density, $J_C$ (kA/cm²)

Josephson Junction Cross Section

Resistively Shunted Junction
MIT-LL Process Progression

SFQ-3ee (4-Metal-Layer, 500nm min features)

- Contact pad
- Vias
- Resistor
- Josephson junction

SFQ-4ee (8-Metal-Layers, 500nm min features)

- Junction Layers
- Wiring Layers

SFQ-5ee (8-Metal-Layers, 350nm min features)

- High-K Layer
- mΩ resistor
Outline

• Short history of SFQ fabrication and circuit development
• State-of-the-art SFQ fabrication
• Scaling limitations and EDA needs
• Summary
Hypres 6-Metal Layer Process (Ripple-2)

- 2 planarized layers beneath standard 4 metal layer process
  - 0.1 kA/cm², 4.5 kA/cm² and 10 kA/cm² Jc processes available
  - Minimum wiring feature size
    - 1 µm for 4-layer process
    - 0.25 µm for 4+-layer process
  - Minimum junction size: 1 µm
- Working to add planarization to junction layers

Ripple-2 Cross Section

- 150-mm wafers, significant refresh in 2012
- ~2500 ft² of class 100/1000
- 248nm stepper
- Planarization using “Caldera” process (etch + CM) for geometry independent planariation
9 Nb layers, M1-M7 planarized
• Additional CMP and thick metal under JJ layer
• 10kA/cm² and 20kA/cm² Jc processes
• Minimum feature size: 1µm
• Minimum junction size: 1µm
• Mo resistors with 2.6 Ω/sq
• Circuits demonstrated 69k JJ (shift register) and 19k JJ (logic)

75-mm wafer tool set
• 2900 ft², class 100 / 1000 cleanroom
• i-line stepper
D-Wave Systems 6-Metal Layer Process
Fabricated at Cypress

• Fully planar 6-metal layer process
• Process Jc is 0.4kA/cm² (for qubits)
• Minimum feature size: 250nm
• Minimum junction size: ~500nm
• Circuits demonstrated with 125k JJs

Cypress foundry in Bloomington, MN

• 200-mm production tool set
• 80k ft² class 10 cleanroom
• 90nm-350nm baseline flows in production
• Development access to production environment
• DMEA trusted foundry
• ~ 400 employees
MIT-LL 8 Metal Layer Process

- Fully planarized 4, 8, and 10 Nb wiring layers
- Minimum feature size 350 nm (248 nm lithography)
- Minimum junction size <500nm
- 10kA/cm², 20kA/cm² and 50kA/cm² (experimental) processes
- Demonstrated circuits with ~70k JJs (shift registers)

MIT-LL Fabrication Facility

- 200-mm production tool set
- 8k ft² class 10 cleanroom
- 90nm baseline flows in prototype circuit quantities
- Multi-use facility: CMOS, Si imagers, Nb SFQ, GaN on Si, MEMS, microfluidics
- DMEA trusted foundry
- ~ 65 dedicated staff; 24/5 operation
# IARPA SFQ Technology Roadmap
*(Government Foundry, MIT-LL)*

<table>
<thead>
<tr>
<th>Fabrication Process Attribute</th>
<th>Units</th>
<th>Process Node</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SFQ3ee SFQ4ee SFQ5ee SFQ6ee SFQ7ee SFQ8ee</td>
</tr>
<tr>
<td>Critical current density</td>
<td>MA/m²</td>
<td>100 100 100 100 100 100</td>
</tr>
<tr>
<td>JJ diameter (surround)</td>
<td>nm</td>
<td>700 (500) 700 (500) 700 (300) 700 (300) 500 (200) 500 (200)</td>
</tr>
<tr>
<td>Nb metal layers</td>
<td>-</td>
<td>4 8 8 10 10 10</td>
</tr>
<tr>
<td>Line width (space)</td>
<td></td>
<td>Critical layers Other layers</td>
</tr>
<tr>
<td></td>
<td>nm</td>
<td>500 (1000) 500 (700) 350 (500) 350 (500) 250 (300) 180 (220) 500 (700) 500 (700) 350 (500) 250 (300)</td>
</tr>
<tr>
<td>Metal thickness</td>
<td>nm</td>
<td>200 200 200 200 200 150</td>
</tr>
<tr>
<td>Dielectric thickness</td>
<td>nm</td>
<td>200 200 200 200 200 180</td>
</tr>
<tr>
<td>Resistor width (space)</td>
<td>nm</td>
<td>1000 (2000) 500 (700) 500 (700) 500 (700) 500 (500) 350 (350)</td>
</tr>
<tr>
<td>Shunt resistor value</td>
<td>Ω/sq</td>
<td>2 2 2 or 6 2 or 6 2 or 6 2 or 6</td>
</tr>
<tr>
<td>mΩ resistor</td>
<td>mΩ</td>
<td>- - 3 - 10 3 - 10 3 - 10 3 - 10</td>
</tr>
<tr>
<td>High kinetic inductance layer</td>
<td>pH/sq</td>
<td>- - 8 8 8 8</td>
</tr>
<tr>
<td>Via diameter (surround)</td>
<td>nm</td>
<td>700 (500) 700 (500) 500 (350) 500 (350) 350 (250) 350 (200)</td>
</tr>
<tr>
<td>Via type, stacking</td>
<td>-</td>
<td>Etched, Staggered Etched, Stacked (\frac{1}{2}) Etched, Stacked (\frac{1}{2}) Etched, Stacked (\frac{1}{2}) Stud/Plug, Stacked Stud/Plug, Stacked</td>
</tr>
<tr>
<td>Early access availability</td>
<td>-</td>
<td>2014 2015 2016 2016 2017</td>
</tr>
</tbody>
</table>

- Drives increased density
- Recently added
MIT-LL SFQ4ee Process

**Process Features**
- Wafer size: 200 mm
- JJ technology: Nb/Al-AlO_x/Nb
- $J_c$: 10 kA/cm^2 (100 µA/µm^2) baseline
- Number of Nb layers: 8
- Min JJ size: 700 nm
- Min wiring size: 500 nm
- Min spacing: 700 nm
- Full planarization of all layers by CMP
- Fab cycle time: <2.5 months, 8 wafers

**Integration Scale Demonstrated**
- AC-biased SFQ shift registers with 32.8k JJs (Semenov, 2015)
- RQL shift registers with 32.8k and 40k JJs on a chip (Herr, 2015)
- AC-biased SFQ shift registers with 65,000+ JJs per circuit (Semenov, SBU)
- AC-biased SFQ shift registers with 144,000+ JJs - under test (Semenov, SBU)
Enhancements over SFQ4ee:

- Min linewidth and spacing: 0.35 μm except M0, M1, and R5 (0.5 μm)
- Min size of vias and metal surround: 0.5 μm and 0.35 μm, respectively
- 2 Ω/sq (Mo) or 6 Ω/sq (MoNₓ) resistors for JJ shunting and biasing
- High kinetic inductance to enable compact bias inductors for ERSFQ
- mΩ resistor between Nb layers M4 and M5
More Deeply Scaled Resistors
MoN$_x$ Resistors

- **Issue:** shunt resistors occupy a considerably larger area than JJs

- In SFQ4ee, resistors made from Mo
  - $w = 0.5 \ \mu m$, $R_s = 2 \ \Omega/sq$, $2A_{\text{via}} \sim 1 \ \mu m^2$
  - RQL circuits use JJs with $I_c \sim 35 \mu A \rightarrow R \sim 20 \ \Omega$, and
    - $A_R \sim 5 \ \mu m^2 \gg A_J$ - JJ area

- In SFQ5ee, option for MoN resistors
  - $R_s = 6 \ \Omega/sq$, $\ell_{\text{min}} = 1.2 \ \mu m$
  - $R = 20 \ \Omega$, $\sim 1 \mu m^2$
  - $T_c < 2 \ K$
  - Wafer $1\sigma = 2.3\%$ on patterned resistors
More Deeply Scaled Bias Circuitry
MoNₓ High-Kinetic-Inductance Layer

- **Issue:** Energy-efficient RSFQ requires multiple 100-pH bias inductors
  - Conventional (geometric) inductor occupies typically ~ 100 μm² area, \( L_g < 1 \) pH/sq

- Kinetic inductance \( L_k \) of thin superconducting films >> geometric inductance:
  - Need film thickness, \( d \), to be much thinner than penetration depth, \( \lambda \)

- For practicality: \( d \sim 35 \text{ nm} - 40 \text{ nm} \), so \( \lambda \) needs to be ~ 400 nm – 500 nm,
  - MoNX with higher nitrogen content is an effective choice
  - Area savings: at \( L_k = 10 \) pH/sq and \( w = 0.7 \) μm, \( A_k \sim A_g/20 \)
Flux Trapping in More Deeply Scaled Circuits
mΩ-range Resistors

- Issue: SFQ circuits are sensitive to flux trapping: external and internal
  - With circuit density increasing, flux trapping could increase
    - Diminishing distance between flux-trapping moats and inductors
    - Diminishing size of the moats
    - Some SFQ cells are particularly sensitive, cannot be reset

- Possible remedies:
  - Break some of the superconducting loops by mΩ-range resistors
  - Improved circuit and moat design

- Fab solution: add extra resistive layer
  - Resistor thickness required to be in the ‘green’ area is between 150 nm and 180 nm
  - Pursuing Mo and MoNₓ resistors
• Stacked vias permit
  – Decrease in size
  – Increase in performance

• Have developed a stud via process that can be stacked
  – Nb-Al-Nb trilayer (without oxide)
  – Diameters down to 250nm

• Significant increase in complexity of the fabrication
  – 2X increase in metal dep, photo masks, CMP time (30% increase in number of steps)
  – Complicates CMP: more stringent density requirements, requires higher uniformity

• In parallel, pursuing a damascene, CVD process for plug or plug+line formation
Process Control Monitor (PCM) Test Data

- PCM reticle included on all SFQ wafers
  - 3255 test structures on 16 chips
  - 9 “drop-out” reticle shots per wafer
- Complete room temperature testing
  - Junction Jc’s, ‘spreads’ and topography effects
  - Via strings
  - Snakes and combs on wiring layers
  - Resistors
- Selected cold temperature testing
  - Junction Jc
  - Inductor structures
  - Line and via critical current
- Extensive software
  - Automated measurement
  - Calculating parameters of interest
  - Assessing process splits and historic comparisons
Advanced Process Test Vehicle
Flux Shuttle Shift Register

• Best to test actual digital circuits
  - Need a digital circuit scalable to ~1M JJs
  - Measure margins of individual cells, identify defects and trapped flux
• Employed a very old (before SFQ) idea: ac-biased, inhomogeneous flux shuttle*
  - 4 JJs per bit
  - Stony Brook U design (V. Semenov)

• Results
  - Can observe trapped flux and operation at cell level
  - Cell operation variation: $1\sigma = 1\%$
  - Exceptional uniformity of Josephson junctions, $1\sigma_{JJ} = 0.8\%$

<table>
<thead>
<tr>
<th>Gen</th>
<th>#Bits</th>
<th>#JJs</th>
<th>Min Line ($\mu$m)</th>
<th>Cell Size ($\mu m^2$)</th>
<th>JJ Density (per cm²)</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>8k</td>
<td>33k</td>
<td>0.5</td>
<td>40x17</td>
<td>0.6M</td>
<td>Func</td>
</tr>
<tr>
<td>2nd</td>
<td>16.4k</td>
<td>66k</td>
<td>0.4</td>
<td>20x15</td>
<td>1.3M</td>
<td>Func</td>
</tr>
<tr>
<td>3rd</td>
<td>36k</td>
<td>144k</td>
<td>0.4</td>
<td>20x15</td>
<td>1.3M</td>
<td>Test</td>
</tr>
<tr>
<td>4th</td>
<td>80.2k</td>
<td>321k</td>
<td>0.4</td>
<td>15x10</td>
<td>2.7M</td>
<td>Fab</td>
</tr>
</tbody>
</table>

* G.M. Lapir, (1977)

4th Gen Flux Shuttle Layout (15x10µm)
Outline

- Short history of SFQ fabrication and circuit development
- State-of-the-art SFQ fabrication
- Scaling limitations and EDA needs
- Summary
Beyond CMOS - 26
MAG 4/4/2016

Circuit Density Based on Josephson Junction Limitations

- \( n_J = k / A_J \), fill factor \( k = 0.5 \)
- Max density limited by the area of RSJ (~ 12 \( \mu \text{m}^2 \) per RSJ)
- 10x increase if get rid of shunts \( \rightarrow \) self-shunted JJs
Circuit yield is determined by:
- Operating margins designed into the circuit
- Variation of the Ic’s in the JJs (Lower Ic’s have larger variations)

Intersection of lines show minimum Ic needed as a function of # JJs and design margin.
Circuit Density Based on Inductor Limitations

- Circuit density is also limited by inductors
  - Each JJ requires an inductor, value is dependent on Ic of JJ

- Red and blue curves show potential inductor density as a function of
  - Inductors on 1 or 2 layers
  - Damping parameter of JJ

- Need to choose Ic greater than intersection of the inductor and JJ density curves
Design Tool Limitations for SFQ Circuits

Electronic design automation (EDA) tools needed at all levels of SFQ design

- **Technology CAD (TCAD)**
  - Junction barrier design
  - Fabrication impact on component performance and tolerance

- **Custom design flows: I-V based circuit simulation**
  - Compact models of Josephson junctions and other components
  - Simulator for circuits with 100s of JJs (J-Spice)
  - Analog circuit design

- **Standard-cell design flow: HDL based design**
  - Standard cells based on different design approaches
  - Timing and data synchronization

- **Physical verification**
  - Layout versus schematic (LVS)
  - Magnetic design rules (flux trapping mitigation)

---

**Standard Cell Design Flow**
(from SuperTools Proposer’s Day)
MIT-LL CMOS Validated Design Flow

- Example EDA flow that supports custom and standard cell designs
SFQ Circuit Hypothetical Design Flow

- Modules requiring substantial modification are highlighted in red
- Maintain modularity and interface approaches common in CMOS EDA
Outline

• Short history of SFQ fabrication and circuit development
• State-of-the-art SFQ fabrication
• Scaling limitations and EDA needs
• Summary
SFQ Fab is as Cheap as Dumbed-Down CMOS

- Current IARPA programs to mature SFQ design and fabrication are vital to the continued maturation of the technology

- IARPA C3 SFQ roadmap
  - On track to develop circuits with JJ density $> 1\text{M/cm}^2$
  - On track to demonstrate circuits with $> 1\text{ MM} \text{ JJs}$
  - 100X increase over state-of-the-art at program start

- With 3rd generation energy-efficient design techniques, limits to the circuit density of $< 10\text{M JJs/cm}^2$
  - SFQ fab could be as cheap as dumbed-down Si processing
  - Can use multi-chip module approaches for near and mid-term demonstrations
SFQ Circuit Technology is Immature: Needs Continual Assessment of Goals

- SFQ has unique attributes that make it relevant for beyond CMOS applications
  - Potential for Landauer limit operation
  - Potential for > 50 GHz clock speed
  - Built-in quantized resources for mixed-signal applications

- Still relatively immature
  - Impressive demonstrations that validate potential of the technology
  - IARPA C3 program can solidify viability of the technology
  - Further investments needed for
    - Develop ‘industrial scale’ EDA tools
    - Maturation of fabrication and testing infrastructure

- Potential paths beyond the C3 horizon
  - Maturation of adiabatic quantum flux parametron
  - Design approaches that are more tolerant of fabrication variations
  - Novel device and circuit approaches for more deeply scaled nodes