

# National Strategic Computing Initiative

## Seminar Program

June 7, 2016 – Feb 7, 2017



**NSF Stampede**  
**U of Texas/Austin**

**DoE Titan/ORNL**

**NSF Blue Waters/NCSA**



**NIST**  
National Institute of  
Standards and Technology  
U.S. Department of Commerce

# NSCI

**On July 29 Obama signed an EO creating the National Strategic Computing Initiative. In order to maximize the benefits of HPC for economic competitiveness and scientific discovery, the United States Government must create a coordinated Federal strategy in HPC research, development, and deployment. Investment in HPC has contributed substantially to national economic prosperity and rapidly accelerated scientific discovery. Creating and deploying technology at the leading edge is vital to advancing my Administration's priorities and spurring innovation. Accordingly, this order establishes the National Strategic Computing Initiative (NSCI).**

# NSCI

- **Started as whole of govt effort and now is a whole of nation effort.**
- **DoE, DoD and NSF are lead agencies**
- **NIST and IARPA are foundational research agencies**
- **NASA, FBI, NIH, DHS and NOAA are deployment agencies**
  - **Each chartered to support the NSCI in ways consistent with the mission and traditional strengths of that agency but to work together to ensure that the US moves forward and to avoid duplication**
- **Importance of NSCI**
  - **Scientific discovery and economic competitiveness required action**
- **Many complex scientific problems rely on both simulation and/or the manipulation, curation and analysis of very large data sets.**
- **Separations of these two aspects are being blurred and new approaches are required if we are to move toward from petascale to exascale.**

- While NIST will play a central role as a foundational agency, there was a clear need both internally and externally to bring current expertise to educate and also to help formulate the path to achieving the NSCI goals.
- NIST decided to start an NSCI seminar program.

### NSCI Internal Web Site

**Organizing Group: Vladimir Aksyuk, Albert Davydov, Bob Hickernell, Curt Richter and Barry I. Schneider**

**Support: Hoyt Cox (AV), Zulma Lainez (G'berg) and Kimberley Keinath (Boulder)**

- The initial thinking was that it would be only an internal seminar program but as a consequence of much broader interest, it was decided to allow outsiders to attend as visitors, subject to the usual NIST rules and to webcast the talks to selected sites. I am happy to say that this has been accomplished.

**Kickoff Talk:**

# **The IARPA Cryogenic Computing Technology program**

**Marc Manheimer**

**Program Director IARPA  
Superconducting Computer Program.**

OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE



# Superconducting Computing



Marc Manheimer  
IARPA Program Manager  
7 June 2016  
NIST NSCI Seminar



# Superconducting Computing

- Cryogenic Computing Complexity Program Motivation
- Program Snapshot
- Technical Background
- Why this is hard
- Miscellaneous



# C3 Program Motivation





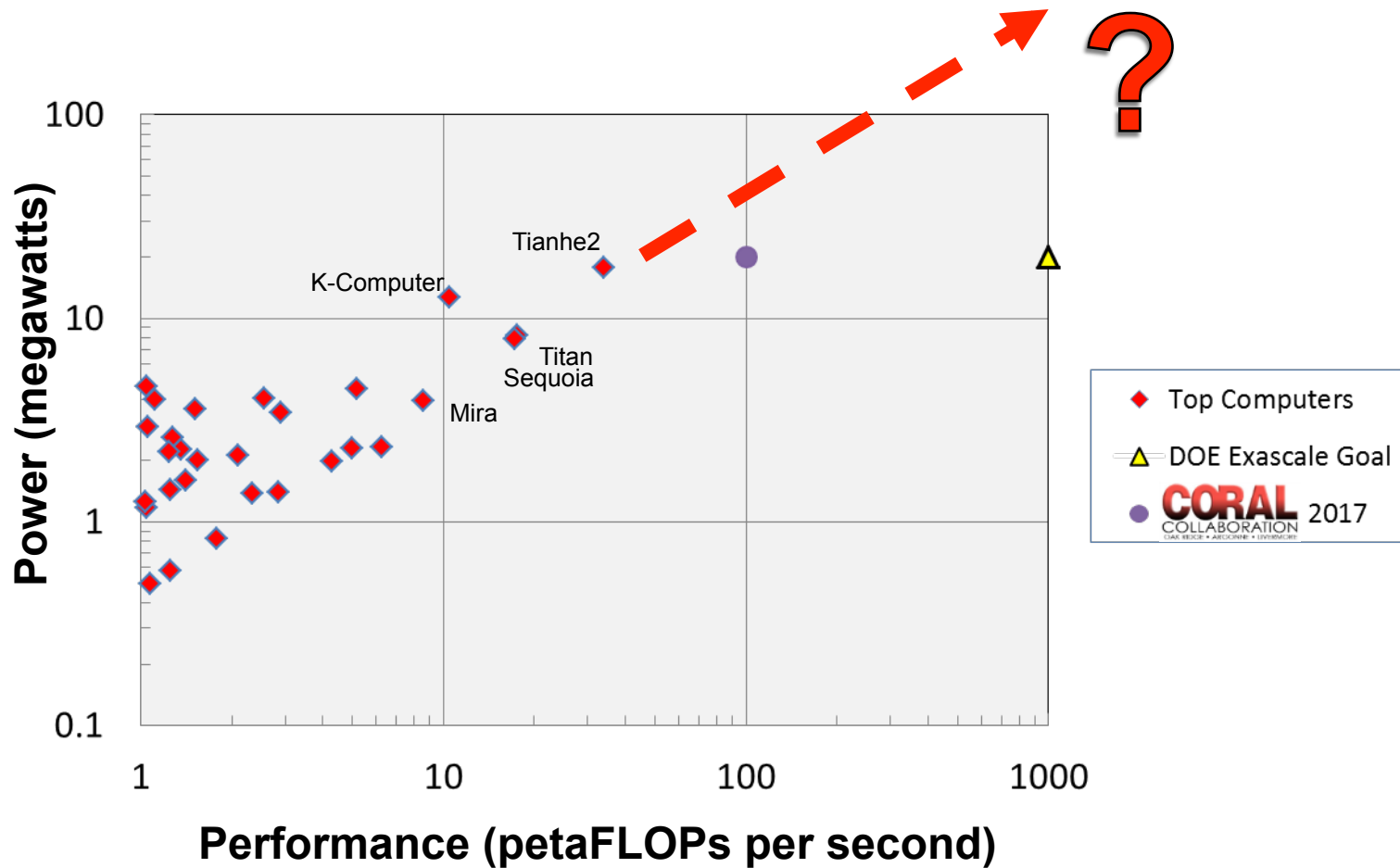
## Our appetite for computing is unlimited!

- Upgrading a facility to more powerful computers is constrained by
  - **Power** supply capability of electric company
  - **Space** limitations
  - **Cooling** infrastructure
- Constraints on developing computers with additional processing power
  - Some estimates to reach exascale are in the hundreds of megawatts.
  - An exascale computer at 20 megawatts based on semiconducting technology will require heroic measures.
  - We will require a different technology to get beyond exascale.

***A computer based on **superconducting logic** and **cryogenic memory** can help solve these issues***

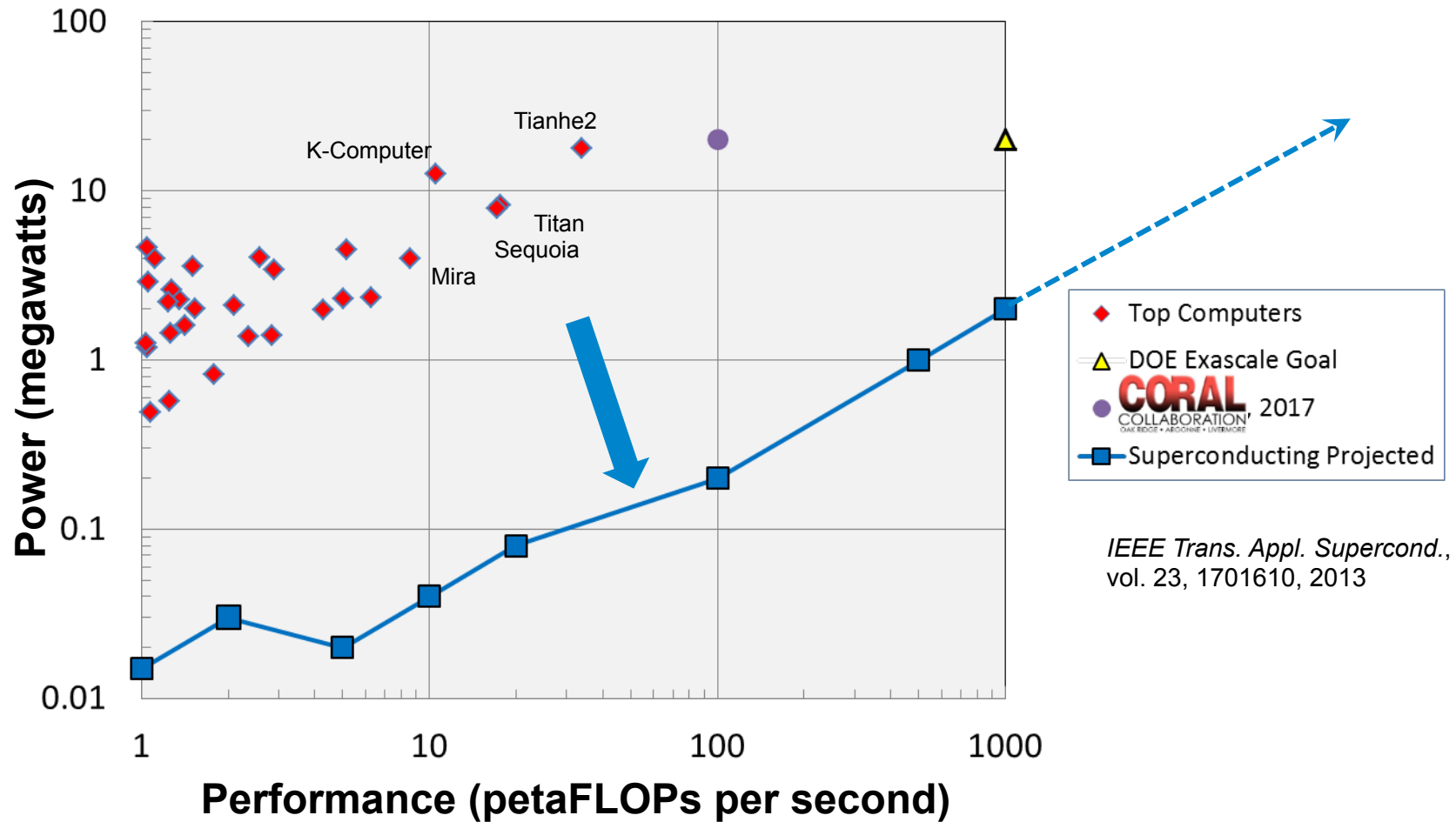


# Problem: Increasing Power Requirements For Conventional Supercomputers





# Superconducting computing looks promising





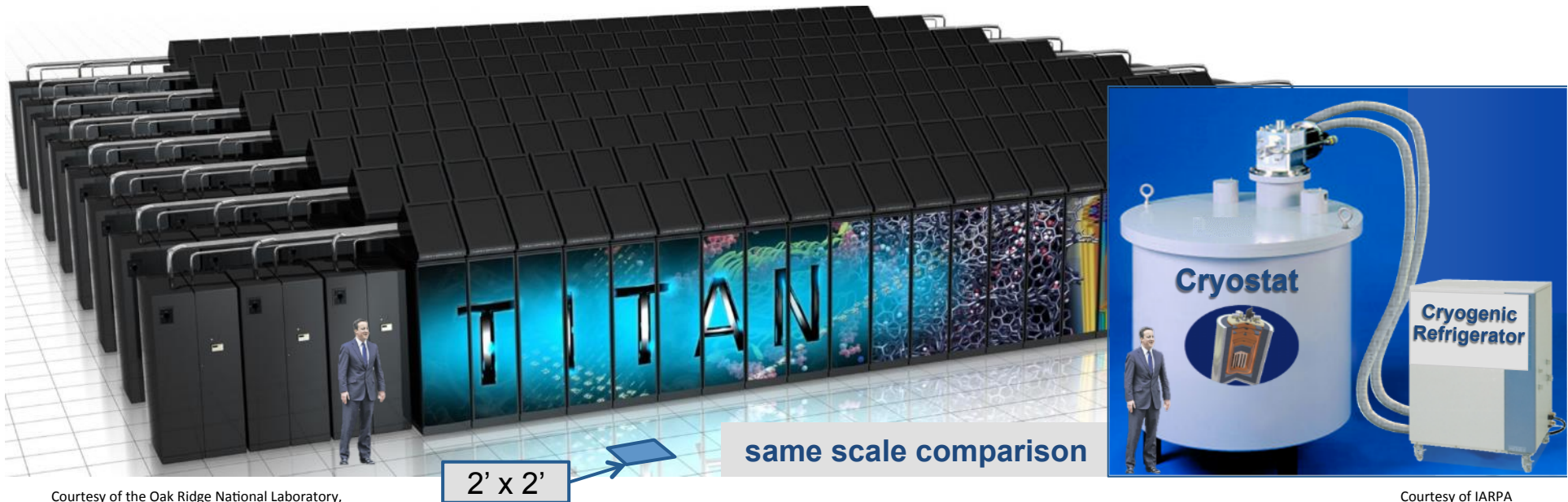
## Key Factors

- Approach based on:
  - ***Near-zero energy*** superconducting interconnect
  - ***New*** SFQ logic with no static power dissipation
  - ***New energy efficient*** cryogenic memory ideas
  - ***Electrical or optical*** ingress/egress
  - ***Commercial*** cryogenic refrigerators

**IARPA C3  
program basis**



# Conceptual System Comparison (~20 PFLOP/s)



Courtesy of the Oak Ridge National Laboratory, U.S. Department of Energy

Courtesy of IARPA

	Titan at ORNL	Superconducting Supercomputer	
<b>Performance</b>	17.6 PFLOP/s (#2 in world*)	20 PFLOP/s	~1x
<b>Memory</b>	710 TB (0.04 B/FLOPS)	5 PB (0.25 B/FLOPS)	7x
<b>Power</b>	8,200 kW avg. (not included: cooling, storage memory)	80 kW total power (includes cooling)	0.01x
<b>Space</b>	4,350 ft <sup>2</sup> (404 m <sup>2</sup> , not including cooling)	~200 ft <sup>2</sup> (19 m <sup>2</sup> , includes cooling)	0.05x
<b>Cooling</b>	additional power, space and infrastructure required	All cooling shown	

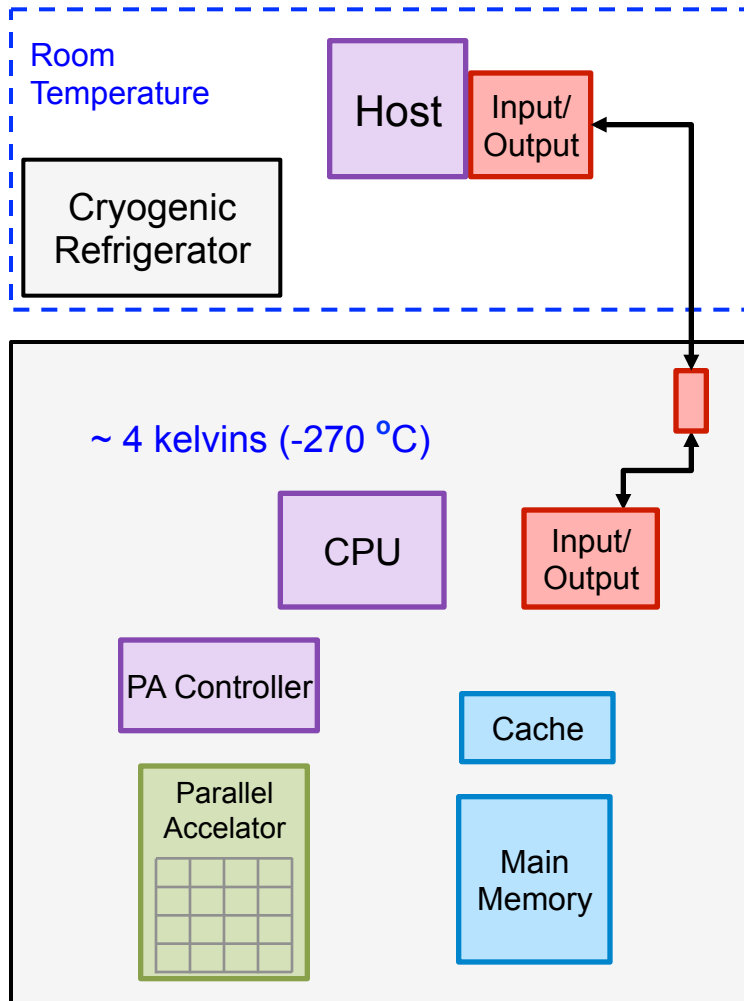
\* #1 in TOP500, 2012-11 (17.6 PFLOP/s)



# C3 Program Snapshot



# C3 Notional System, Metrics, and Goals



Metric	Goal
Clock rate for superconducting logic	10 GHz
Throughput (bit-op/s)	$10^{13}$
Efficiency @ 4 K (bit-op/J)	$10^{15}$
CPU count	1
Word size (bit)	64
Parallel Accelerator count	2
Main Memory (B)	$2^{28}$
Input/Output (bit/s)	$10^9$



## C3 Phase 1 Metrics and Goals

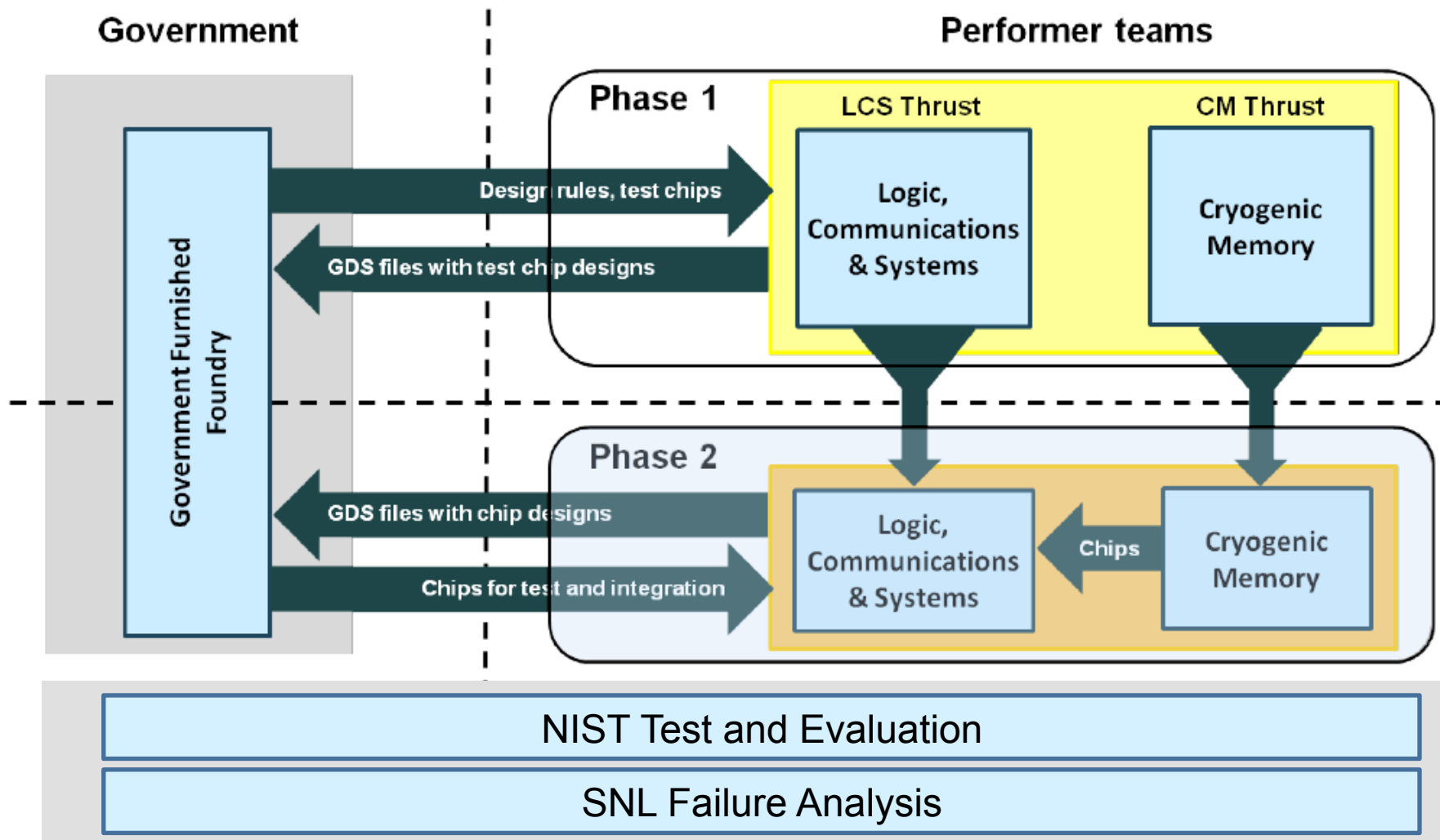
Metric	BP	OP1	OP2
Cryogenic Memory	Memory cell	Array	Chip
Functional capacity (bit)*	1	$2^6$ ; $2^6$	$2^{10}$ ; $2^{10}$
Density (bit/cm <sup>2</sup> )*	$10^6$ ; $10^5$	$5 \times 10^6$ ; $5 \times 10^5$	$10^7$ ; $10^6$
Data rate, burst mode (Gbit/s)*	1	5; 30	5; 30
Access time, ave. (ps)*	10,000; 1,000	5,000; 400	5,000; 400
Access energy, ave. (J/bit)*	$5 \times 10^{-16}$ ; $5 \times 10^{-17}$	$5 \times 10^{-16}$ ; $5 \times 10^{-17}$	$10^{-16}$ ; $10^{-17}$
Logic, Comm. & Systems	Subcircuits	Circuits	Processors
Benchmark circuits & applications	Circuits 1	Circuits 2	Circuits 3
Complexity (JJ)	$10^4$	$5 \times 10^4$	$10^5$
Density (JJ/cm <sup>2</sup> )	$10^5$	$5 \times 10^5$	$10^6$
Throughput (bit-op/s)	$10^9$	$5 \times 10^{10}$	$10^{11}$
Efficiency @ 4 K (bit-op/J)	$10^{16}$	$5 \times 10^{16}$	$10^{17}$

\* Memory metrics: The first number refers to Main Memory and the second to Cache Memory.





# C3 Organization





## C3 Augmentations

- IARPA SuperTools program – BAA is in preparation
- Official Request For Information on data ingress and egress
  - Published on FedBizOpps; responses received in March
  - Likely new program



# Performer Teams

## Cryogenic Memory

***NORTHROP GRUMMAN***

### **Northrop Grumman Corporation**

- Arizona State University
- Michigan State University

### **Raytheon BBN Technologies**

- **Raytheon**  
BBN Technologies
- Cornell University
- Massachusetts Institute of Technology
- New York University
- University of Rochester

## Logic, Communications, & Systems

**IBM**

### **IBM Watson Research Center**

- Hypres Inc.
- Rensselaer Polytechnic Institute
- Stony Brook University

***NORTHROP GRUMMAN***

### **Northrop Grumman Corporation**



# Government Teams

## Foundry

MIT Lincoln Lab



## Test & Evaluation

NIST Boulder

**NIST**  
National Institute of  
Standards and Technology  
U.S. Department of Commerce



JPL



*Contracting in Process*



Sandia National Labs





# Technical Background

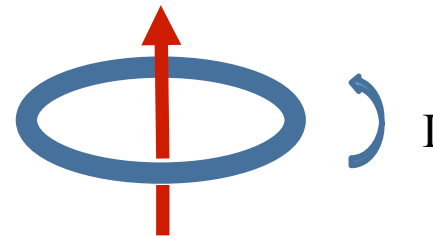


# Superconductors

- Current flows without resistive losses below a critical temperature

- Electrical current in a loop stores flux

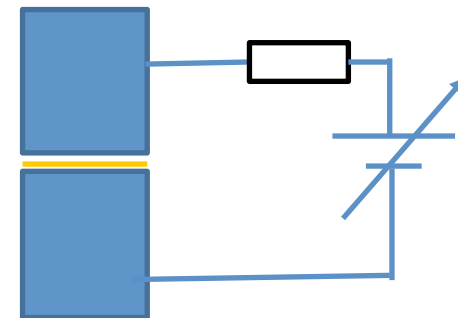
- Can be basis for memory,  $I \bullet L = \Phi$
- Flux is quantized; minimum value is  $\Phi_0 = h/2e \approx 2.07 \text{ fWb (mV}\cdot\text{ps or mA}\cdot\text{pH)}$
- To hold one flux quantum, a 20 micrometer diameter niobium loop would carry about 30 microamperes



- Supercurrent can flow across a thin insulating barrier between two superconductors:

## Josephson tunneling

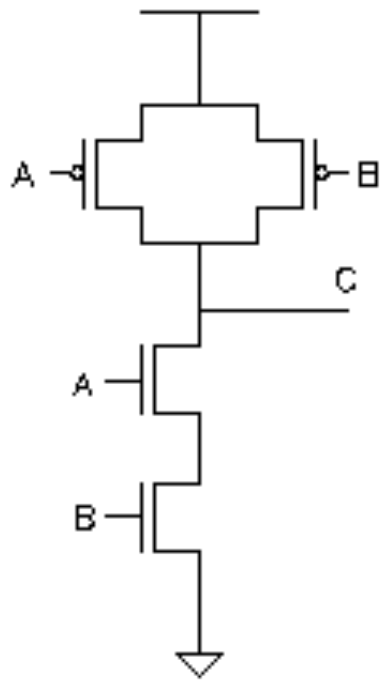
- This works up to a critical current,  $I_c$
- For  $I > I_c$ , an ac voltage develops across the barrier
- For a short input pulse  $\int V dt = \Phi_0$
- Typical pulses are a picosecond long and millivolts tall, with energy  $\sim 10^{-19}$  joules



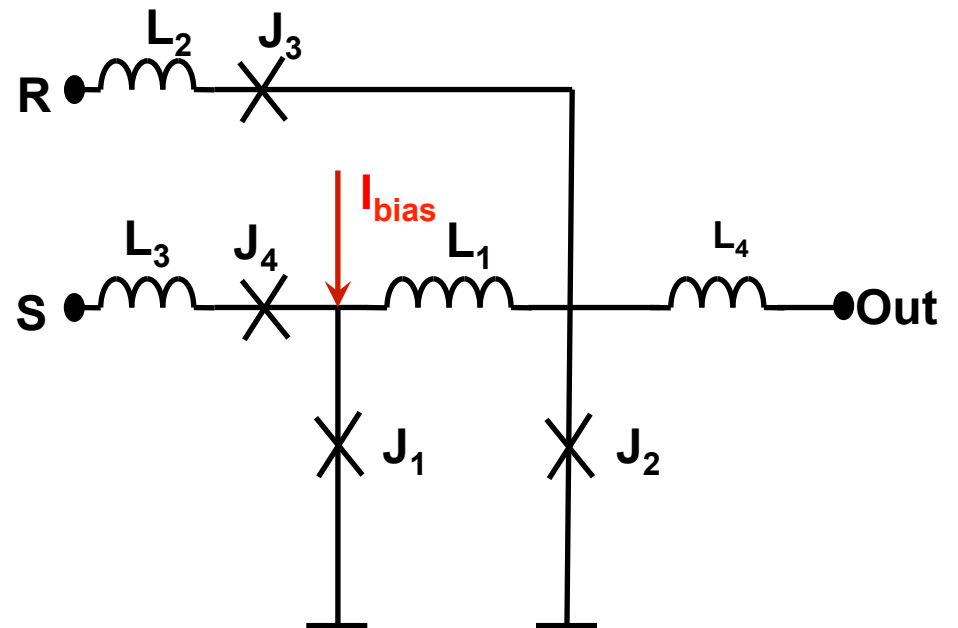




# Simple Circuits: NAND and RS Flip-flop



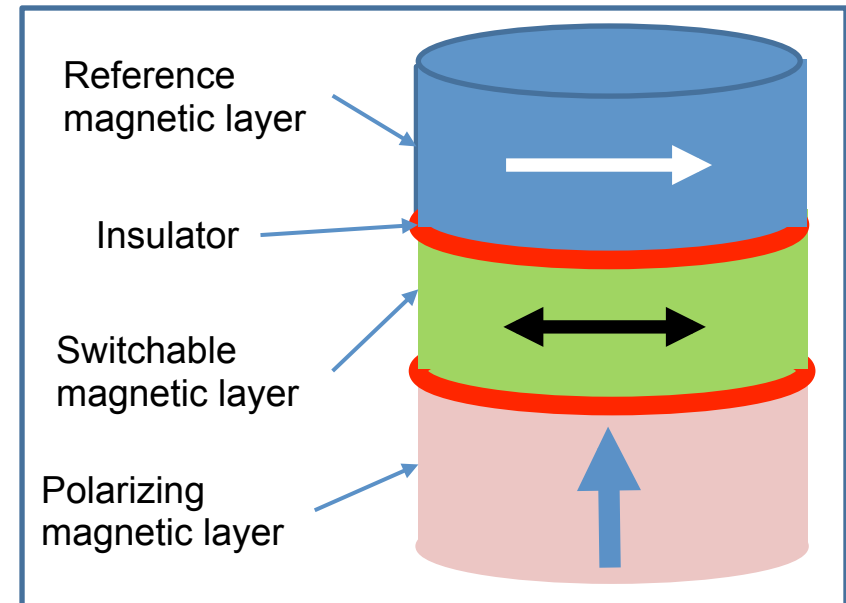
NAND





## COST Memory (Raytheon-BBN, NYU)

- Stands for Cryogenic Orthogonal Spin Transfer
- Requires multiple magnetic layers (MTJ structure)
- Electron current is spin polarized in passing through permanent magnet polarizing layer
- Polarized electrons exert torque on switchable magnetic layer
- Magnetization of switchable layer is reversed
- This changes resistance of stack.

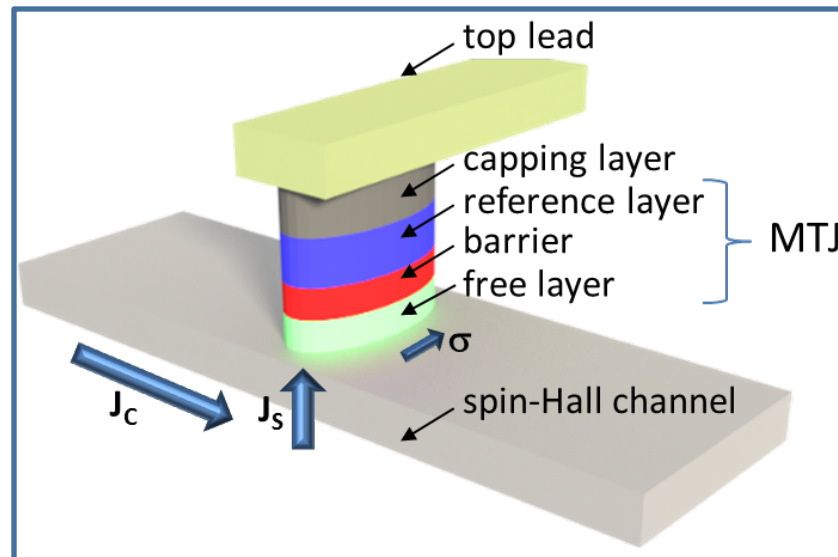






# Cryogenic Spin Hall Effect Memory (Raytheon-BBN, Cornell U)

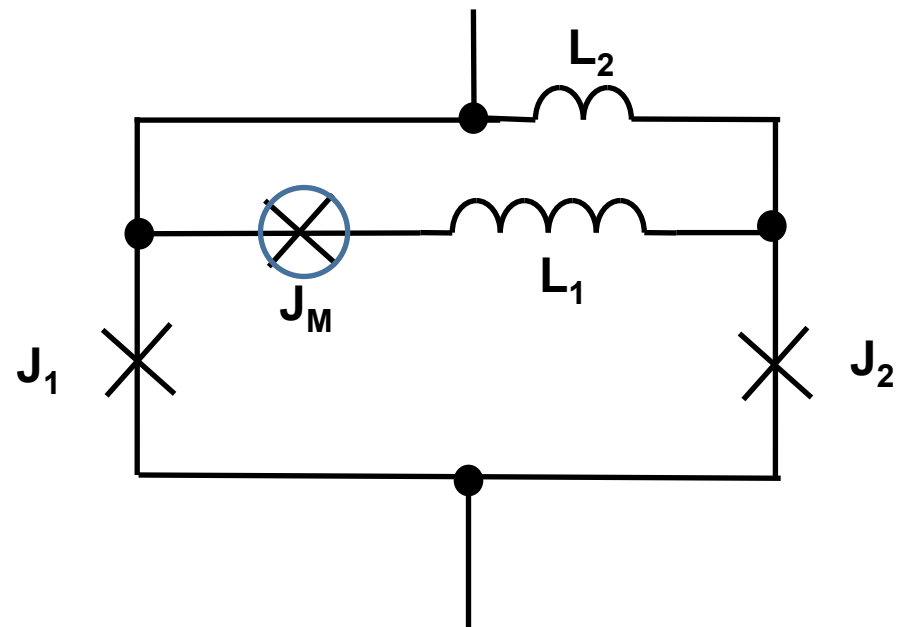
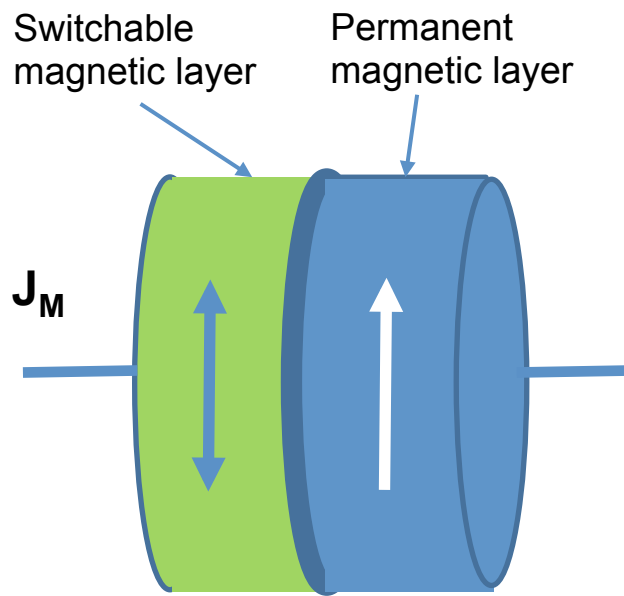
- Spin orbit coupling of the electron current in the spin-Hall channel gives rise to an orthogonal spin current that goes into the magnetic tunnel junction (MTJ) structure.
- The spin current can reverse the magnetization of the free layer.
- Magnetoresistance is read by passing current through the MTJ structure.





## JMRAM Memory (NGC, MSU, ASU)

- Memory cell requires a special kind of magnetic Josephson junction that is switchable so that the superconducting wave function will experience either a zero or pi phase shift.
- With a pi phase shift, the magnetic junction creates a spontaneous current, and the critical current of the circuit is measurably reduced.



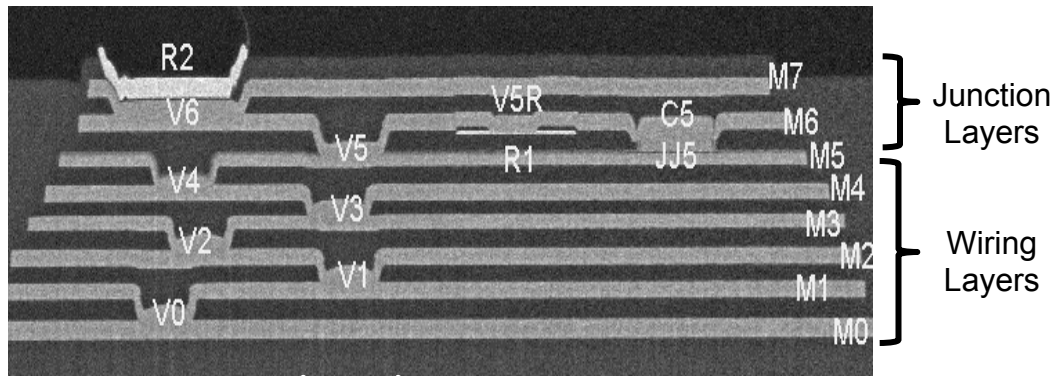


## Why this is hard



# Advanced Fabrication Process

- No commercial foundry exists that can fabricate circuits at the required level of complexity.
  - MIT Lincoln Laboratory (LL) has a niobium superconductor circuit foundry that IARPA is upgrading to meet the aggressive program goals.
  - With 200 mm wafers and 8+ planarized niobium layers, **the MIT LL superconducting foundry is now the most advanced in the world – and continues to advance.**
  - LL will transfer the technology elsewhere as directed.
- ✓ Test chips with up to 144 k JJs all working have been successfully demonstrated in the 8-Nb-layer process. **We are now only four orders of magnitude behind CMOS!**



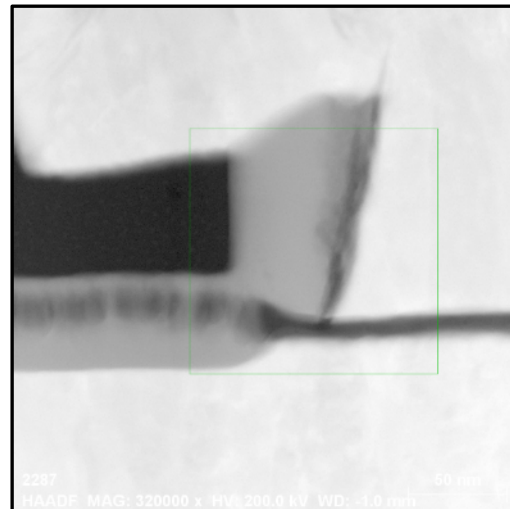
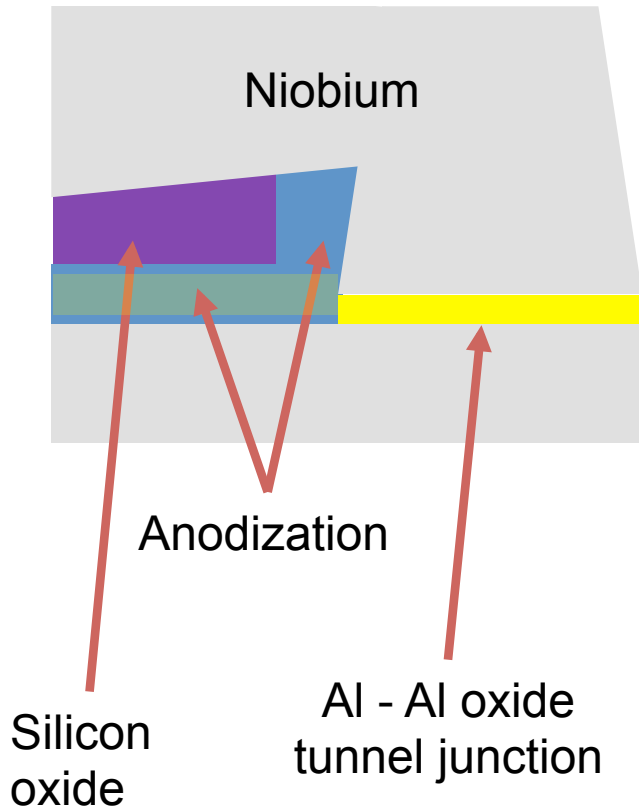
8 niobium layer cross-section



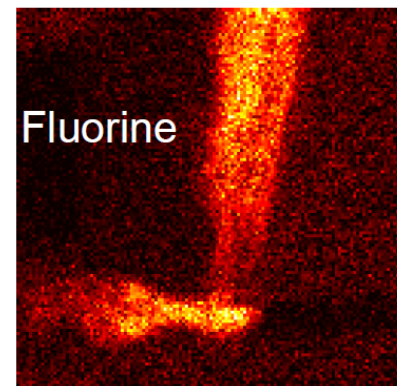
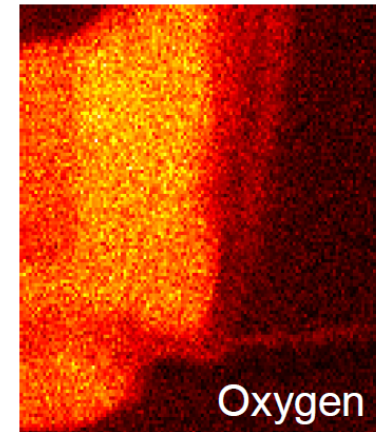


# Lincoln Laboratory Yield Problem

Poor cleaning after fluorine etch may lead to anodization separating from niobium electrode



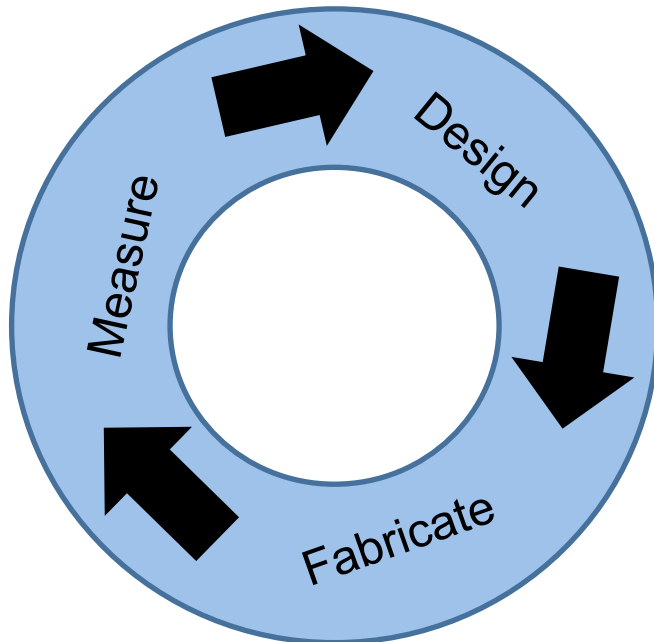
Transmission Electron Microscope Photo from Sandia





# Superconducting Logic Development

- Designing complex circuits without software design tools is like ***making bricks without straw***
  - Some software development is included in C3
  - SuperTools may help by C3 Phase II
- Learning requires multiple fabrication turns



- Complications include
  - Incomplete set of design rules
  - Design rules change for every advance in technology node
  - Multi-project wafer runs are inherently less reliable than single project runs
  - Unanticipated problems with fabrication

- Close communication between design teams and LL fab is essential for success





# Foundry Technology Roadmap

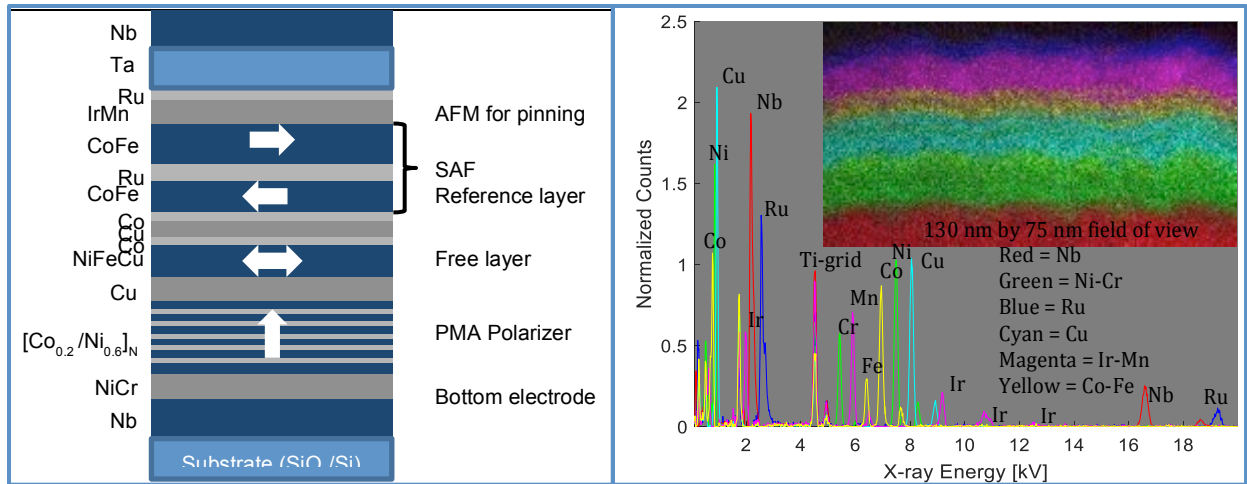
Fabrication Process Attribute	Units	Process Node					
		SFQ3ee	SFQ4ee	SFQ5ee	SFQ6ee	SFQ7ee	SFQ8ee
Critical current density	MA/m <sup>2</sup>	100	100	100	100	100	100
JJ diameter (surround)	nm	700 (500)	700 (500)	700 (300)	700 (300)	500 (200)	500 (200)
Nb metal layers	-	4	8	8	10	10	10
Line width (space)	Critical layers	500 (1000)	500 (700)	350 (500)	350 (500)	250 (300)	180 (220)
	Other layers			500 (700)	500 (700)	350 (500)	250 (300)
Metal thickness	nm	200	200	200	200	200	150
Dielectric thickness	nm	200	200	200	200	200	180
Resistor width (space)	nm	1000 (2000)	500 (700)	500 (700)	500 (700)	500 (500)	350 (350)
Shunt resistor value	Ω/sq	2	2	2 or 6	2 or 6	2 or 6	2 or 6
mΩ resistor	mΩ	-	-	3 - 10	3 - 10	3 - 10	3 - 10
High kinetic inductance layer	pH/sq	-	-	8	8	8	8
Via diameter (surround)	nm	700 (500)	700 (500)	500 (350)	500 (350)	350 (250)	350 (200)
Via type, stacking	-	Etched, Staggered	Etched, Stacked \2/	Etched, Stacked \2/	Etched, Stacked \2/	Stud, Stacked	Stud, Stacked
Early access availability	-		2014-09	2015-09	2016-03	2016-09	2017-09

Future →

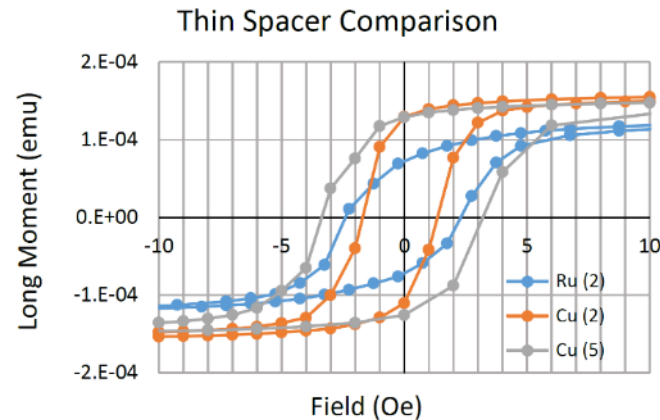
Changes from the previous process

# Magnetic Memory Optimization

- Optimizing magnetic layers
  - Shape, size
  - Thickness
  - Materials
  - Temperature dependence
  - Fabrication
    - Crystalline anisotropy
    - Smoothness
    - Uniformity
    - Hygiene effects



- Decoders
- Drivers
- System requires integration of diverse technologies onto a single substrate



[Nb(25)/Al(2.4)]<sub>3</sub>/Nb(20)/Spacer/NiFe(1.5)/Spacer/Nb(5)



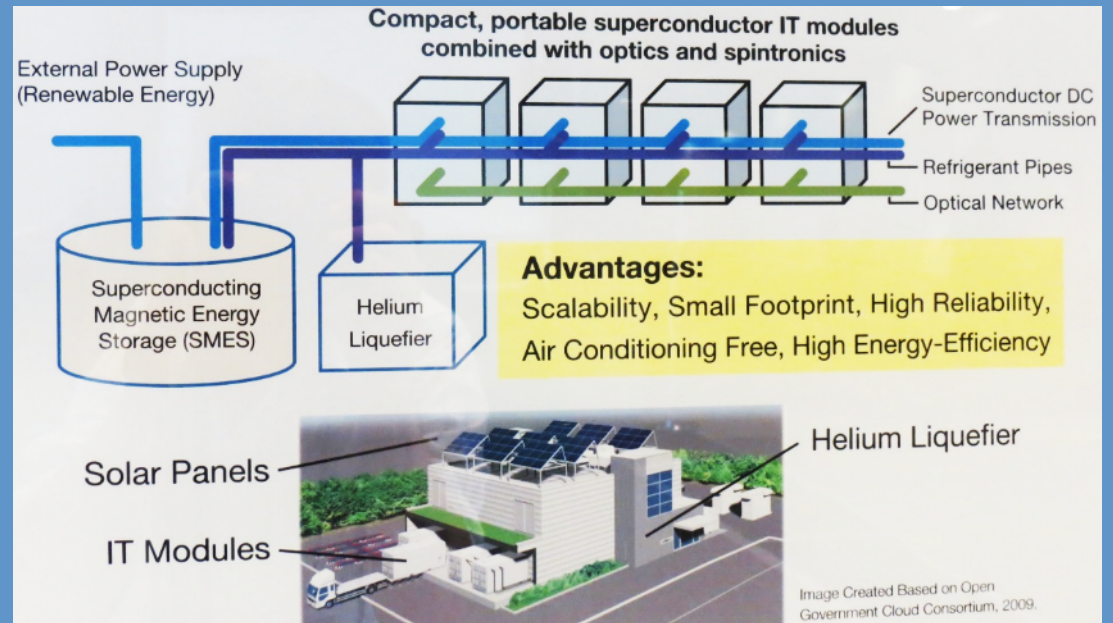
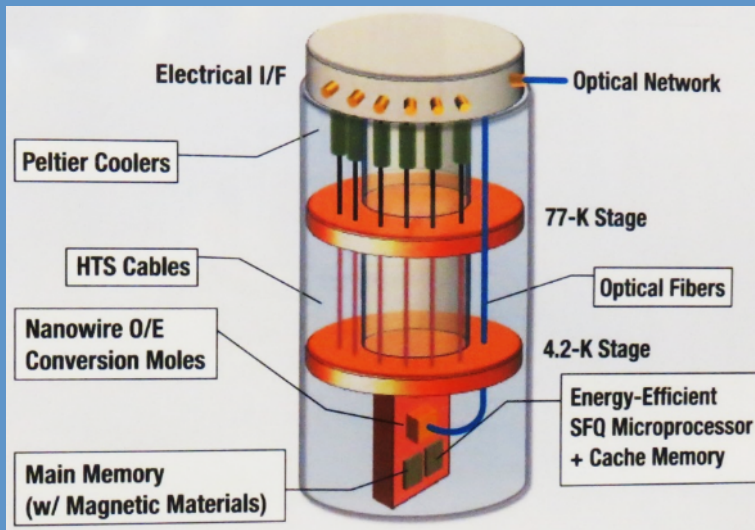


# Miscellaneous



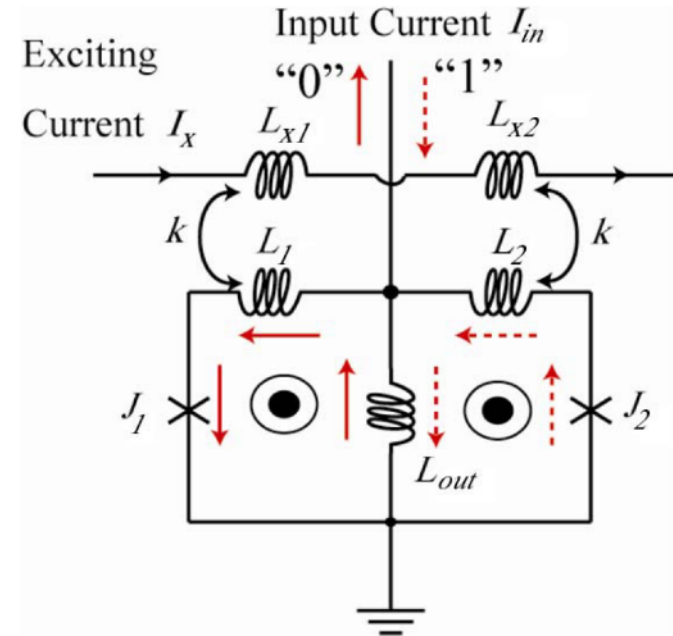
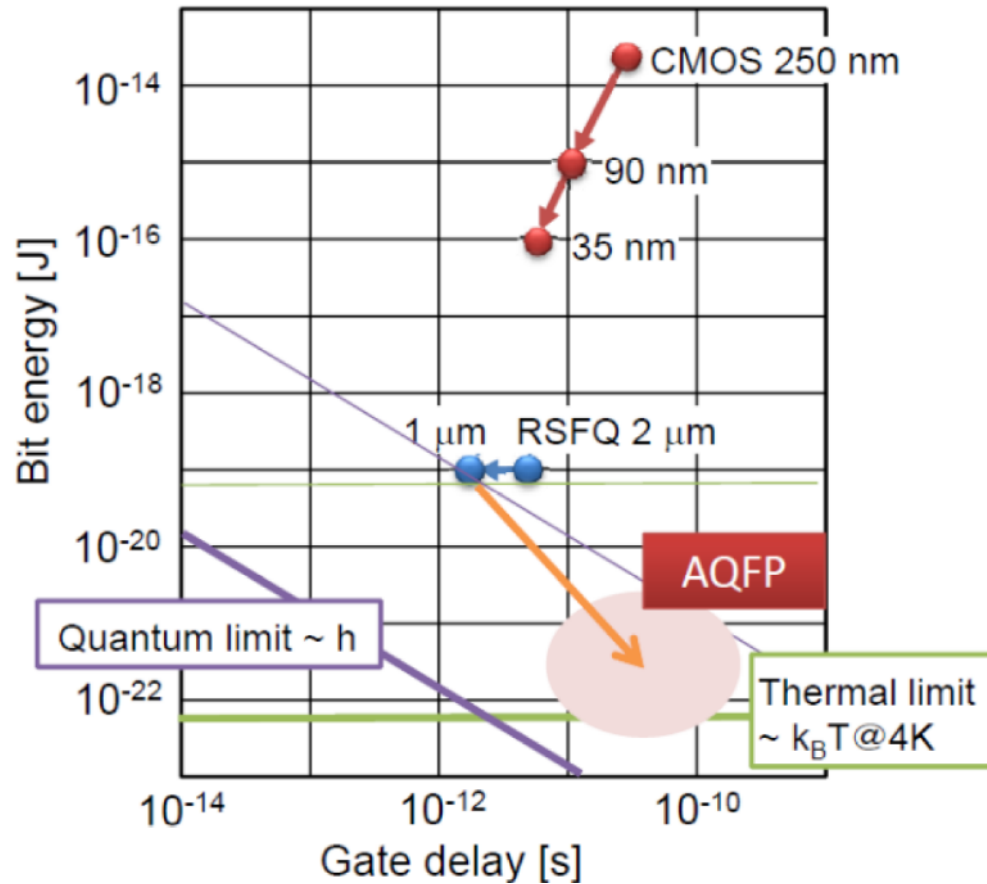
## Japanese 3-year program started

- “Superconductor Electronics System Combined with Optics and Spintronics”  
JST-ALCA Project: <http://www.super.nuqe.nagoya-u.ac.jp/alca/> (Japanese)
- Processor goals: AQFP logic, 8-bit simplified RISC architecture, ~25,000 JJs, ~10 instructions





# Adiabatic Quantum Flux Parametron



K. Inoue, N. Takeuchi, K. Ehara, Y. Yamanashi, and N. Yoshikawa, IEEE TRANSACTIONS ON APPLIED SUPERCONDUCTIVITY, VOL. 23, NO. 3, JUNE 2012

Nobuyuki Yoshikawa, IEEE/CSC & ESAS European Superconductivity News Forum (ESNF) No. 24 April 2013



# New Superconducting Spintronics Programme

Cambridge to research future computing tech that could “ignite a technology field”

Research ^

News ^

[Cambridge to research future computing tech that could “ignite a technology field”](#)



Media enquiries

[Tom Kirk](#)

[Communications office](#)

Published

15 Apr 2016

Image

Growing quantities of data storage online are driving up the energy costs of high-performance computing and data centres. Superconducting spintronics offer a potential means of significantly increasing their energy-efficiency to resolve this problem.

Credit: [10515 images via Pixabay](#)

Share

Email 0 reddit 0

Facebook Share 544 LinkedIn Share 117

Tweet 215 sharethis 881

Facebook Like 1.2K

The project will be led by Professor Mark Blamire, Head of the Department of Materials Sciences at the University of Cambridge, and Dr Jason Robinson, University Lecturer in Materials Sciences, Fellow of St John’s College, University of Cambridge, and University Research Fellow of the Royal Society. They will work with partners in the University’s Cavendish Laboratory (Dr Andrew Ferguson) and at Royal Holloway, London (Professor Matthias Eschrig).

**A Cambridge-led project aiming to develop a new architecture for future computing based on superconducting spintronics - technology designed to increase the energy-efficiency of high-performance computers and data storage - has been announced.**

A project which aims to establish the UK as an international leader in the development of “superconducting spintronics” – technology that could significantly increase the energy-efficiency of data centres and high-performance computing – has been announced.

Led by researchers at the University of Cambridge, the “Superspin” project aims to develop prototype devices that will pave the way for a new generation of ultra-low power supercomputers, capable of processing vast amounts of data, but at a

“*Superconducting spintronics offer extraordinary potential because they combine the properties of two traditionally incompatible fields to enable ultra-low power digital electronics*”

– Jason Robinson



## National Strategic Computing Initiative (NSCI)

### Executive Order July 29, 2015

By the authority vested in me as President by the Constitution and the laws of the United States of America, and to maximize benefits of high-performance computing research, development, and deployment, it is hereby ordered as follows:

- (b) Foundational Research and Development Agencies. **There are two foundational research and development agencies for the NSCI:** the Intelligence Advanced Research Projects Activity (**IARPA**) and the National Institute of Standards and Technology (**NIST**).

**IARPA will focus on future computing paradigms offering an alternative to standard semiconductor computing technologies.** NIST will focus on measurement science to support future computing technologies. The foundational research and development agencies will coordinate with deployment agencies to enable effective transition of research and development efforts that support the wide variety of requirements across the Federal Government.





## NSCI Objectives

1. Exascale computing system (~100x performance relative to present)
2. Increase coherence between the technology base used for modeling and simulation and that used for data analytic computing.
3. Establishing, over the next 15 years, a viable path forward for future HPC systems even after the limits of current semiconductor technology are reached (the "post-Moore's Law era").
4. Increase the capacity and capability of an enduring national HPC ecosystem by employing a holistic approach that addresses relevant factors such as networking technology, workflow, downward scaling, foundational algorithms and software, accessibility, and workforce development.
5. Develop an enduring public-private collaboration to ensure that the benefits of the research and development advances are, to the greatest extent, shared between the United States Government and industrial and academic sectors.



## IARPA NSCI Portfolio

- Cryogenic Computing Complexity
- Quantum Computing Programs
  - QEO: seeks tenfold increase in quantum bit coherence time
  - Multi-Qubit Coherent Operations: seeks to resolve challenges of multiple qubit technology systems
  - LogiQ: seeks to build a better logical qubit out of imperfect physical qubits
- Brain Inspired MICrONS program: seeks to revolutionize machine learning by reverse-engineering algorithms of the brain

**IARPA is seeking new ideas and program managers to fill out its NSCI portfolio**



**Thank you!**