Cryogenic Computing Complexity (C3)

• Marc Manheimer
• marc.manheimer@iarpa.gov
• December 9, 2015
• IEEE Rebooting Computing Summit 4
C3 for the Workshop

• Review of the C3 program
  – Motivation
  – Technical challenges
  – Program details

• The National Strategic Computing Initiative and its impact on C3 and IARPA
The Problem is Power-Space-Cooling

- 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*
C3 Goal:

Develop technologies for a computer based on superconducting logic with cryogenic memory, and integrate a prototype that can answer these questions:

1) Can we build a superconducting computer that is capable of solving important problems?

2) Does it provide a sufficient advantage over conventional computing that we want to build it?
Problem: Increasing Power Requirements For Conventional Supercomputers
Superconducting computing looks promising

![Graph showing performance vs. power for different supercomputers.](image_url)

System Comparison (~20 PFLOP/s)

Titan at ORNL

<table>
<thead>
<tr>
<th>Performance</th>
<th>17.6 PFLOP/s (#2 in world*)</th>
<th>20 PFLOP/s</th>
<th>~1x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>710 TB (0.04 B/FLOPS)</td>
<td>5 PB (0.25 B/FLOPS)</td>
<td>7x</td>
</tr>
<tr>
<td>Power</td>
<td>8,200 kW avg. (not included: cooling, storage memory)</td>
<td>80 kW total power (includes cooling)</td>
<td>0.01x</td>
</tr>
<tr>
<td>Space</td>
<td>4,350 ft² (404 m², not including cooling)</td>
<td>~200 ft² (includes cooling)</td>
<td>0.05x</td>
</tr>
<tr>
<td>Cooling</td>
<td>additional power, space and infrastructure required</td>
<td>All cooling shown</td>
<td></td>
</tr>
</tbody>
</table>

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

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

**IARPA C3 program basis**
Technical Challenges 1

• **Memory:** energy-efficient, fast, dense, useful capacity, compatible with superconducting single flux quantum (SFQ) logic for direct integration
  – C3 ideas include MRAM, spin Hall effect, JMRAM, nMEM
  – Requires interface circuits of significant complexity in SFQ technology
  – Requires understanding new physics with interplay of spintronics and superconductivity

• **Logic complexity:** designing superconducting integrated circuits with far more elements on a single chip than previously achieved
  – In SFQ computing, the devices are Josephson junctions and the logic elements are picosecond wide pulses; these present new design challenges.
  – Electronic design automation tools are either missing or not scalable to very large scale integration
Technical Challenges 2

- **Advanced fabrication process**: multilayer, sub-micrometer feature size with specialty layers (high kinetic inductance, milliohm resistance ...)
  - Variance of key fabrication parameters ($J_c$, inductance) must be improved
  - Must develop detailed simulation module that includes process variations
  - Close coupling between circuit design-test and process design rules is key
  - Need to develop close coupling between foundry and failure analysis team

- **System**: demonstrate a superconducting computer with multiple processors and memory in MCM packaging; beyond C3 challenges include:
  - Scalable system design
  - Wafer-scale stacking with superconducting through silicon vias
  - High data-rate interconnect between 4 K and room temperature
Advanced Fabrication Process

• A commercial foundry that can fabricate circuits at the required level of complexity does not exist.

• Lincoln Laboratory (LL) has a niobium superconductor circuit foundry that IARPA is upgrading to meet the aggressive program goals

• The superconducting facility at LL is now the most advanced in the world – and is continuing to advance

• LL is working with performers and with potential transition partners to ensure that foundry capability is transferred.
C3 Organization

• Two thrusts:
  – Logic, communications and systems
  – Cryogenic memory

• Two phases:
  – Phase 1, performers develop technology for subsystems
  – Phase 2, performers scale up and integrate technology into a working prototype
Prototype

<table>
<thead>
<tr>
<th>Metric</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock rate for superconducting logic</td>
<td>10 GHz</td>
</tr>
<tr>
<td>Throughput (bit-op/s)</td>
<td>$10^{13}$</td>
</tr>
<tr>
<td>Efficiency @ 4 K (bit-op/J)</td>
<td>$10^{15}$</td>
</tr>
<tr>
<td>CPU count</td>
<td>1</td>
</tr>
<tr>
<td>Word size (bit)</td>
<td>64</td>
</tr>
<tr>
<td>Parallel Accelerator count</td>
<td>2</td>
</tr>
<tr>
<td>Main Memory (B)</td>
<td>$2^{28}$</td>
</tr>
<tr>
<td>Input/Output (bit/s)</td>
<td>$10^9$</td>
</tr>
</tbody>
</table>
Program Metrics and Goals

<table>
<thead>
<tr>
<th>Metric</th>
<th>BP</th>
<th>OP1</th>
<th>OP2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cryogenic Memory</strong></td>
<td>Memory cell</td>
<td>Array</td>
<td>Chip</td>
</tr>
<tr>
<td><strong>Functional capacity (bit)</strong></td>
<td>1</td>
<td>2^6; 2^6</td>
<td>2^{10}; 2^{10}</td>
</tr>
<tr>
<td><strong>Density (bit/cm^2)</strong></td>
<td>10^6; 10^5</td>
<td>5\times10^6; 5\times10^5</td>
<td>10^7; 10^6</td>
</tr>
<tr>
<td><strong>Data rate, burst mode (Gbit/s)</strong></td>
<td>1</td>
<td>5; 30</td>
<td>5; 30</td>
</tr>
<tr>
<td><strong>Access time, ave. (ps)</strong></td>
<td>10,000; 1,000</td>
<td>5,000; 400</td>
<td>5,000; 400</td>
</tr>
<tr>
<td><strong>Access energy, ave. (J/bit)</strong></td>
<td>5\times10^{-16}; 5\times10^{-17}</td>
<td>5\times10^{-16}; 5\times10^{-17}</td>
<td>10^{-16}; 10^{-17}</td>
</tr>
<tr>
<td><strong>Logic, Comm. &amp; Systems</strong></td>
<td>Subcircuits</td>
<td>Circuits</td>
<td>Processors</td>
</tr>
<tr>
<td><strong>Benchmark circuits &amp; applications</strong></td>
<td>Circuits 1</td>
<td>Circuits 2</td>
<td>Circuits 3</td>
</tr>
<tr>
<td><strong>Complexity (JJ)</strong></td>
<td>10^4</td>
<td>5\times10^4</td>
<td>10^5</td>
</tr>
<tr>
<td><strong>Density (JJ/cm^2)</strong></td>
<td>10^5</td>
<td>5\times10^5</td>
<td>10^6</td>
</tr>
<tr>
<td><strong>Throughput (bit-op/s)</strong></td>
<td>10^9</td>
<td>5\times10^{10}</td>
<td>10^{11}</td>
</tr>
<tr>
<td><strong>Efficiency @ 4 K (bit-op/J)</strong></td>
<td>10^{16}</td>
<td>5\times10^{16}</td>
<td>10^{17}</td>
</tr>
</tbody>
</table>

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

- Program is approaching the one year mark

- Two logic performers
  - IBM
  - Northrop Grumman

- Two memory performers
  - Raytheon BBN
  - Northrop Grumman
C3 Government Team

• NIST-Boulder
  – Provides expert technical advice
  – Test and evaluation of performer circuits

• Sandia National Laboratories
  – Provides advanced failure analysis tools
  – Includes superconductivity expertise

• NASA-JPL
  – Optimizing engineering of aluminum oxide tunnel barrier
  – Investigating use of alternate barriers
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 (floating point) and that used for data analytic (integer) 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.
NSCI Roles

- Lead Agencies
  - Department of Energy (DOE)
  - Department of Defense (DOD)
  - National Science Foundation (NSF)

- Foundational Research and Development Agencies
  - IARPA
  - NIST

- Deployment Agencies
  - National Aeronautics and Space Administration,
  - Federal Bureau of Investigation,
  - National Institutes of Health,
  - Department of Homeland Security
  - National Oceanic and Atmospheric Administration

- Executive Council
  - co-chaired by: Director of the Office of Science and Technology Policy (OSTP), Director of the Office of Management and Budget (OMB)
IARPA and NSCI

• IARPA will attempt to fill technology gaps in the current C3 program by
  – New design tool development program
  – New room temperature to cryogenic temperature interconnect program
  – We are currently working on internal IARPA superconductivity-related strategy; this will likely encompass an application space broader than C3.

• IARPA is seeking to develop a broad portfolio of ‘beyond Moore’ computing-related projects
  – New program ideas
  – New program managers

Send me ideas – superconductivity related and not.