Summary Report for the

SRC SPRING 2021 TECH FORUM ON

2030 SEMICONDUCTOR AGENDA

Meeting Dates:	May 4-6, 2021
Meeting Location:	Virtual event

Challenge

The current hardware-software paradigm in information and communication technologies (ICT), manifested by Moore's Law, is reaching its limits and must change. The Industry needs a new plan that would replace the previous roadmap established decades earlier. In this plan, it is important to identify significant trends that are driving information and communication technologies as well as the roadblocks and opportunities that industry faces.

Scientific advances in semiconductors are crucial to economic growth. In fact, information, along with energy, have been the social-economic growth engine for civilization since its earliest beginnings. Arguably, semiconductors have been the largest enabler of today's digital information society; therefore, the limits of Moore's law and 2-dimensional scaling is worrisome. Our society is hungry for information. We need more and more bits and bytes to produce, store, and communicate all that data. *Thus, foundations for the new semiconductor era must be laid down.* As a first step, Semiconductor Research Corporation (SRC) and its partner, the Semiconductor Industry Association (SIA), has launched a new industry-wide road-mapping initiative called the 2030 Decadal Plan for Semiconductors (https://www.src.org/about/decadal-plan/). Therein, the Decadal Plan Committee identified five seismic shifts to attack through innovative research.

The purpose of this Forum was to drive the conversion of the high-level Grand Goals of the Decadal Plan into a detailed *Semiconductor Agenda* toward 2030, driven by SRC. This forum summary is intended to provide input for planning and funding decisions regarding future research programs.

I. New compute trajectories for energy-efficiency

Seismic Shift
Ever-rising energy demands for computing versus global energy production are creating new risk, but new computing paradigms offer opportunities with dramatically improved energy efficiency.

Grand Goal
The Computing Grand Goal is to discover computing paradigms/architectures with a radically new computing trajectory, demonstrating >1,000,000x improvement in energy efficiency. Changing the trajectory not only provides immediate improvements but also provides many decades of growth potential.

Contributors: Hal Finkel (DOE/SC ASCR), John E. Kelly III (IBM), Edlyn Levine (MITRE Corp.), Margaret Martonosi (NSF & Princeton), Partha Ranganathan (Google), Michael Schulte (AMD), Gilroy Vandentop (Intel), Shimeng Yu (GA Tech)

Current Landscape.

Every major advance in humanity in the last 50 years has been enabled by semiconductors. Information and Communication Technologies (ICT) is a direct driver of economic growth, social well-being, and equality. But, today, we are at a crossroads and in need of revolutionary new computing paradigms that ensure sustainable growth for ICT. With a collaborative research agenda, we can rise to meet the needs of our moment and ensure future generations are positioned to thrive.

In addition to the Seismic Shift in computation identified by the Decadal Plan, the current landscape of computing can be characterized by the following observations:

(i) Since 2018 machines started to generate more data than humanity
(ii) There has been an explosion in system complexity

A panel of experts has reviewed the Decadal Plan findings and developed specific recommendations to the Semiconductor Agenda toward 2030 as outlined in the following section.

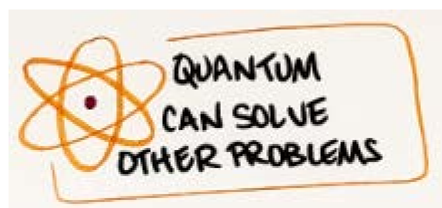
Input to the semiconductor agenda priorities

1. *Verifiability and security research is key along with reliability and resiliency*. We must build this up front into the complex systems of the future. It is not a workable/responsible approach to build systems and tack these on later.

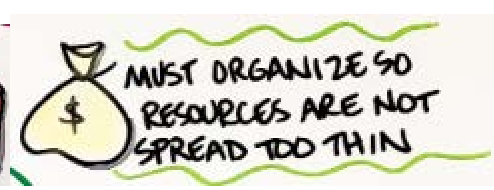
2. *Reconfigurability in complex compute systems.* It is recognized that specialized compute systems can yield energy efficiency, and the resultant heterogeneity creates the need for *reconfigurability*. The growing system complexity drives a need for abstraction to enable simplicity of programming. There is a tension between simplicity of programming and code level optimization for peak energy efficiency. This drives the need for research into machine programming and AI systems which enable the complex verifiability, security, and energy efficiency goals to be addressed simultaneously, which is a task beyond normal human capabilities.

3. *Interfaces and Connectivity Research* is needed to develop system designs that minimizes data movement. Compute in Memory is an example of promising research vector. <u>The brain shows us it is possible and we need to keep learning to understand how</u>.

4. *Quantum Computing Research*. When Quantum Computing is realized it will help us solve currently unsolvable compute problems. Quantum Computers will not supplant traditional computing. We must work on both.



5. Call for Action: There is a known gap in our ability to transfer academic breakthroughs into commercially viable technology solutions. A big roadblock within this gap is academic access to commercially relevant equipment infrastructure to enable the maturing of research ideas and tech transfer. *A National Semiconductor Technology Center should be carefully designed* to address this gap. This is an expensive endeavor but a fair price to pay compared to the price of failure.



II. Future analog electronics and intelligent sensing

Seismic Shift
Fundamental breakthroughs in analog hardware are required to generate smarter world-machine interfaces that can sense, perceive, and reason.

Grand Goal
The Analog Grand Goal is for revolutionary technologies to increase actionable information with less energy, enabling efficient and timely (low latency) sensing-to-analog-to-information with a practical reduction ration of 10^5:1.

<u>Contributors:</u> Ahmad Bahai (Texas Instruments), Ali Keshavarzi (DARPA), Chiao Liu (Facebook), Rikky Muller (UC Berkeley), Boris Murmann (Stanford), Dave Robertson (Analog Devices), Usha Varshney (NSF), Jim Wieser (Texas Instruments)

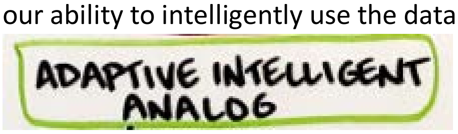
Current Landscape.

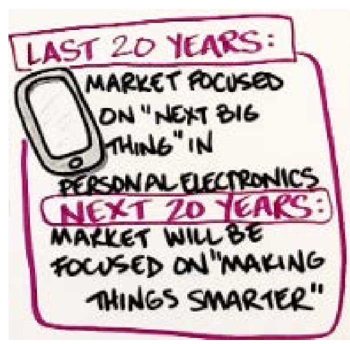
Today, it is impossible to imagine daily life without analog microchips as they connect us to the external world. Analog chips produce the sounds of speech and music in our smartphones, measure our body temperature in electronic thermometers, and bring images to life on our computers, TVs, and other devices. Automobiles and houses contain hundreds of analog chips that enable everything from Bluetooth to sensors to GPS and more.

The current landscape of analog electronics can be characterized by the following observations:

(i) The analog data deluge: Since 2017 the total analog information generated from the physical world by semiconductor sensors has surpassed the collective human sensing throughput.

(ii) Today, the ability to generate analog data is growing faster than our ability to intelligently use the data.



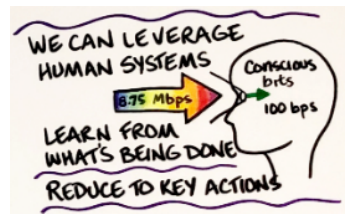


A panel of experts has reviewed the Decadal Plan findings and developed specific recommendations to the Semiconductor Agenda toward 2030 as outlined in the following section.

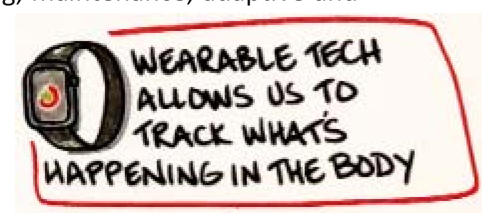
Input to the semiconductor agenda priorities

1. Key areas of research to drive the Grand Goal of 100,000:1 data reduction
    a. Holistic Sensing System Optimization – multi-domain, multi-dimension, multi-level
        i. Architecture and System design methodologies/tools– facilitate partition and block exploration
        ii. Heterogeneous technology application, partitioning, and integration
        iii. Multi-disciplinary/domain design approach understanding "needed" performance (sensor, analog, digital, software, power), and not more than needed.

    b. Sensor technology providing significant "information" versus raw data – including multimodal sensor fusion
    c. Specialized, highly efficient sensor processing to action including power management and actuation
        i. Local processing for sensing to action (closed loop) – robust, secure, meets application need
        ii. Best technology for the domain and processing effectiveness w/minimal data movement
        iii. New analog devices for sensing, interface, and processing

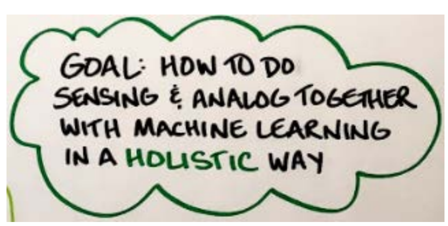


2. Key applications research driving value and addressing the Grand Goal – Sensing to Action focus
    a. Health, Medical/Medicine & Human Interface – monitoring, maintenance, adaptive and active feedback
        i. Drives constrained optimization: power/energy, size, cost, autonomy–multiple applications & interest to young researchers



    b. Infrastructure – dynamic/adaptive/learning multimodal continuous monitoring - Climate, Manufacturing, Energy, Building
    c. Local autonomous/adaptive systems – robotics, autonomous vehicles (ground and air)
        i. Independent of need for human intervention
    d. Virtual Reality/Augmented Reality

3. Metrics to achieve the *data reduction* Grand Goal? (may be sector or application specific)
    a. System value: autonomy, power efficiency, robustness/accuracy, latency, self-powered – NOT individual block FOM (Figure of Merit)
    b. Speed of development: productivity and time from research to commercial product



III. New communication trajectories

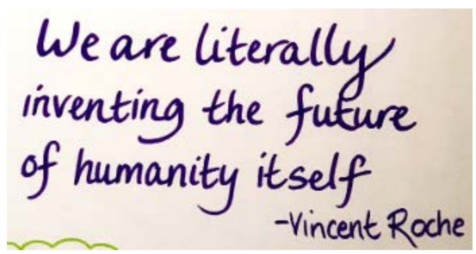
Seismic Shift
Always available communication requires new research directions that address the imbalance of communication capacity vs. data generation rates.

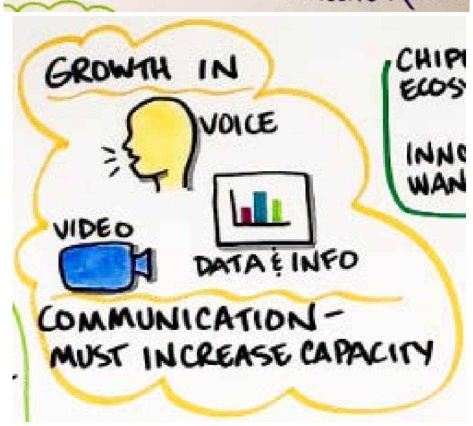
Grand Goals
The Communication Grand Goals include:
- Advance communication technologies to enable moving around all stored data of 100-1000 zettabyte/year at the peak rate of 1Tbps@<0.1nJ/bit
- Develop intelligent and agile networks that effectively utilize bandwidth to maximize network capacity

Contributors: Vincent Roche (Analog Devices), Ramesh Chauhan (Qualcomm), Marla Dowell (NIST), Daniel Friedman (IBM), Sachin Katti (Stanford), Lawrence Loh (MediaTek), Dipankar Raychaudhuri (Rutgers), James Wilson (DARPA)

Current Landscape.

Electronic communication technologies play a critical role in modern world and influence all aspects of our life. In fact, the telecommunication services have spiked in 2020-21, and they become instrumental to minimize the impact of the global health and economic crisis due to the COVID-19 pandemic. There is major infrastructure behind all of it and benefiting from it: the communication industry and all of its components from network providers to cloud systems, to mobile apps and ICT security. This is a massive area of the economy, which is accelerating 5G deployment and future 6G. Telecommunication is becoming a major source of innovation, for example it will be enabler of the Fourth Industrial Revolution (or Industry 4.0), which will be characterized by a massive deployment of IoT devices with seamless connectivity (both wired and wireless).





In addition to the Seismic Shift in communication identified by the Decadal Plan, the current landscape of communication technologies can be characterized by the following observations:

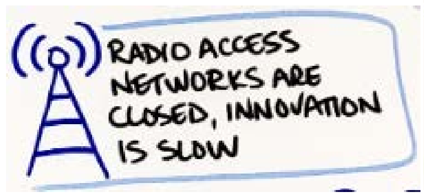
(i) mmWave yields a dramatic increase in achievable throughput
(ii) Massive IoT: >350B connected intelligent devices by 2030
(iii) Increasing role of AI/ML for channel and network optimization
(iv) Trend: Open radio access networks (O-RAN)

A panel of experts has reviewed the Decadal Plan findings and developed specific recommendations to the Semiconductor Agenda toward 2030 as outlined in the following section.

Input to the semiconductor agenda priorities

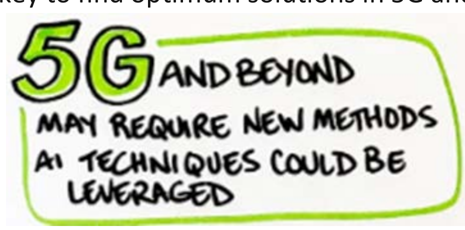
1. Research and effort on multiple fronts are needed to make progress towards the grand goals.
    a. Semiconductor Technology (new devices, 3D, chiplets)
    b. Design, Architectures & Standards (network efficiency, massive MIMO, AI)
    c. Applications

2. Disaggregation in HW (Chiplets) and SW (O-RAN) will be key for innovation and developing agile communication Networks.



3. Software defined Radios will be an essential part of future networks.

4. AI-ML based techniques are key to find optimum solutions in 5G and beyond networks.



5. Current focus is on mmWave based communication systems.

6. Industry-Academia-Government partnership programs would be necessary to deal with challenges of communication grand goals.

IV. Memory and Storage after 2030

Seismic Shift
The growth of memory demands will outstrip global silicon supply, presenting opportunities for radically new memory and storage solutions.

Grand Goals

- **Memory Grand Goal:** Develop emerging memories and memory fabrics with >10-100X density and energy-efficiency improvement for each level of the memory hierarchy
- **Storage Grand Goal:** Discover storage technologies with >100x storage-density capability and new storage systems that can leverage these new technologies

Contributors: James Ang (PNNL), Krste Asanovic (UC Berkeley), Eric Cheng (DOD), Balint Fleischer (Micron), Maya Gokhale (LLNL), Jin Lim (SK hynix), Sayeef Salahuddin (UC Berkeley).

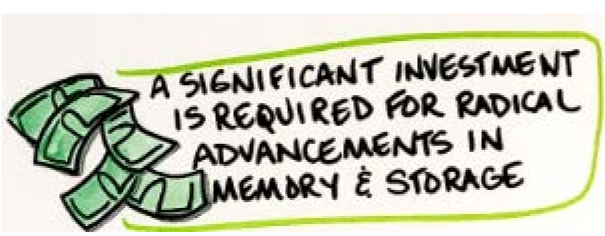
Current Landscape.

Memory and storage are playing increasingly important role in modern ICT and efforts to increase memory/storage capacity are underway. However, the device technology innovations cannot progress without simultaneous development of new interfaces. In this next decade, many technology interfaces are in flux, and it is important to define a new set of interfaces to avoid the need for every vertically integrated computing solution be a custom one-off special purpose design.

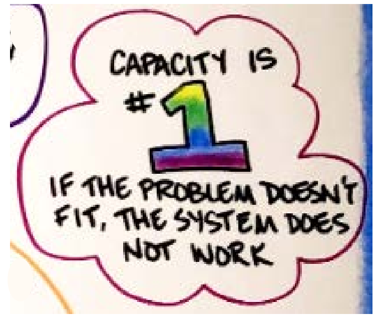
In addition to the Seismic Shift in communication identified by the Decadal Plan, the current landscape of communication technologies can be characterized by the following observations:

(i) Expansion of applications for memory devices (e.g., IoT, edge computing, cloud data center etc.)
(ii) Increasing capacity without equivalent increase in bandwidth translates to massive memory wall bottleneck
(iii) AI-oriented solutions put new demands for memory

A panel of experts has reviewed the Decadal Plan findings and developed specific recommendations to the Semiconductor Agenda toward 2030 as outlined in the following section.



Input to the semiconductor agenda priorities



1. Capacity is #1 factor: if the problem does not fit the system, it does not work!

2. *Intelligent Memory* – New memory architectures to support big data processing such as Process near Memory (PNM), Process in Memory (PIM) and Computing in Memory (CIM)

3. High-bandwidth low latency I/Os, such as photonic

4. New memory hierarchy

5. High-density embedded memory

6. Government investments are instrumental for new memory/storage device technology research

V. Security and Privacy

Seismic Shift
Breakthroughs in hardware research are needed to address emerging security challenges in highly interconnected systems and artificial intelligence.

Grand Goal
Develop security and privacy advances that keep pace with technology, new threats, and new use cases. Examples include trustworthy and safe autonomous and intelligent systems, secure future hardware platforms, and emerging post-quantum and distributed cryptographic algorithms.

Contributors: Mark Segal (NSA), Debra Delise (Analog Devices), Anand Rajan (Intel), Brett Hamilton (DoD), Hungwen Li (MediaTek), George Coker (NSA), Daphne Yao (VA Tech), Thorsten Holz (Ruhr-University Bochum)

Current Landscape

Today, the field of security and privacy is undergoing rapid transformation as new use cases, new threats, and new platforms emerge.

A panel of experts has reviewed the Decadal Plan and would like to emphasize the following observations:

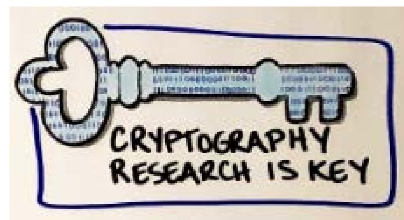
(i) Explosion in system complexity adds to the security challenge
(ii) Pervasiveness of privacy challenges due to IoT
(iii) Existing SW/HW mechanisms cannot defend against Insider and AI attacks

Further, these experts have specific recommendations to the Semiconductor Agenda toward 2030 as outlined in the following section.

Input to the semiconductor agenda priorities

1. *Security research is key along with reliability and resiliency*. We must promote a philosophy of designing in security for the complex systems of the future. It is neither a workable nor responsible approach to design systems and "tack on" security at the end.

2. Chiplets, while providing opportunities to create SoCs with advanced technologies, also create a need for additional security measures.



3. Cryptography research is fundamental.

4. Development and enforcement of formal methods deserves attention.

5. Measurable security is needed.

6. Anticipating the next unknown attacks (attacks not known at design stage) through adaptive response methodologies and system resiliency are increasingly important.

7. Call for Action: Development of workforces that understand the complete hardware/software stack to improve system level security