



THE ELECTRONICS RESURGENCE INITIATIVE

CONSOLIDATED SLIDES AS PRESENTED

LIVESTREAM ARCHIVE AVAILABLE AT
[YOUTUBE.COM/WATCH?V=3hXZR9LTRGA](https://www.youtube.com/watch?v=3hXZR9LTRGA)

DEFENSE APPLICATIONS PROPOSERS DAY:
Leveraging ERI Technologies for Revolutionary Defense Capabilities
Dec 19, 2018 | Arlington, VA

Agenda / Table of Contents

Bookmarks and agenda link to each presentation.

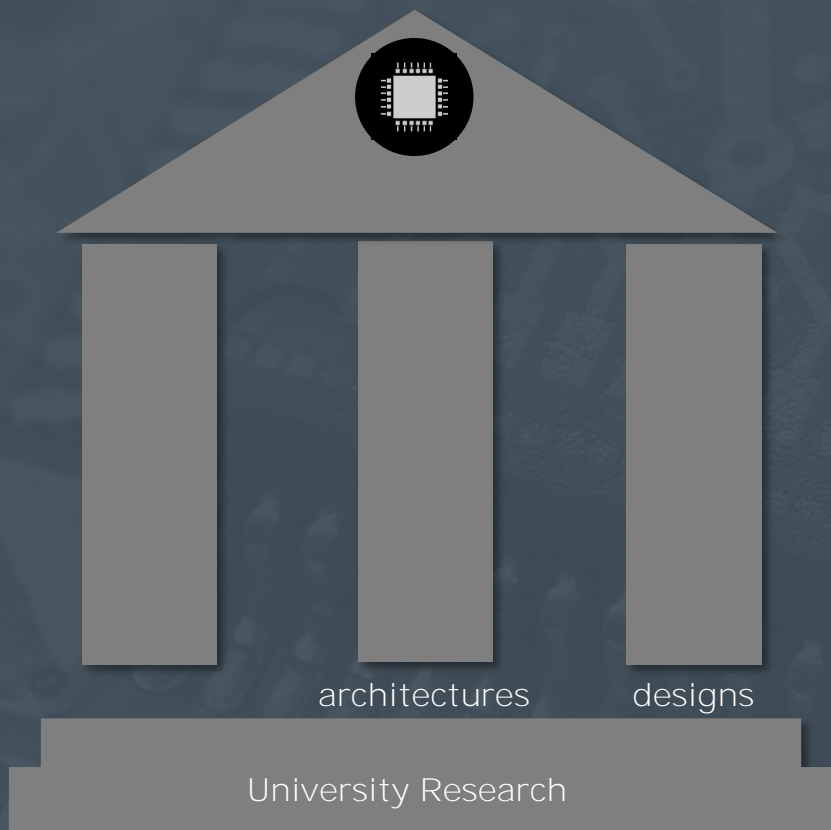
Start	Session		PDF Page
7:45	<i>Check-In (Online registration required before deadline. No onsite registration.)</i>		---
9:00	Greg Woosley	Security Discussion	---
9:05	Bill Chappell	DARPA ERI and Defense Applications	3
9:20	Richard-Duane Chambers	Enabling Collaboration	19
9:35	Michael Blackstone	Contracting Guidance	29
10:05	Wade Shen	SDH	42
10:30	Thomas Rondeau	DSSoC	56
10:55	Collaboration Period / Poster Session		---
11:30	Lunch		---
12:30	Andreas Olofsson	CRAFT, IDEA, POSH, CHIPS	69, 80, 105
13:10	Young-Kai Chen	FRANC	118
13:35	Jay Lewis	3DSoC, JUMP	130, 138
14:00	Collaboration Period / Poster Session		---
14:30	Q&A Period (<i>Notecard submission</i>)		---
15:30	Collaboration Period		---
16:00	END		



THE ELECTRONICS RESURGENCE INITIATIVE

What's Next for ERI?

Dr. William Chappell
Director, Microsystems Technology Office (MTO)



Materials & Integration Thrust

How do we integrate new materials for specialized functions?



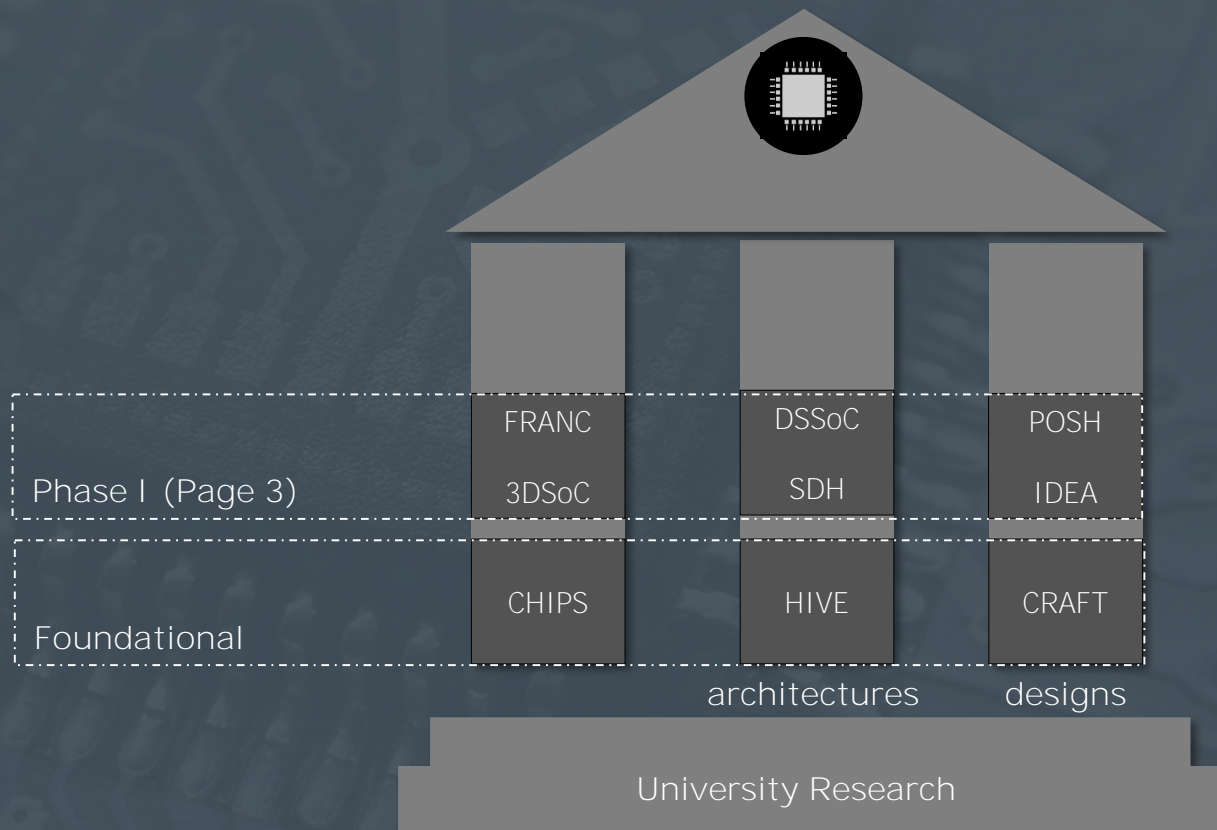
Designs Thrust

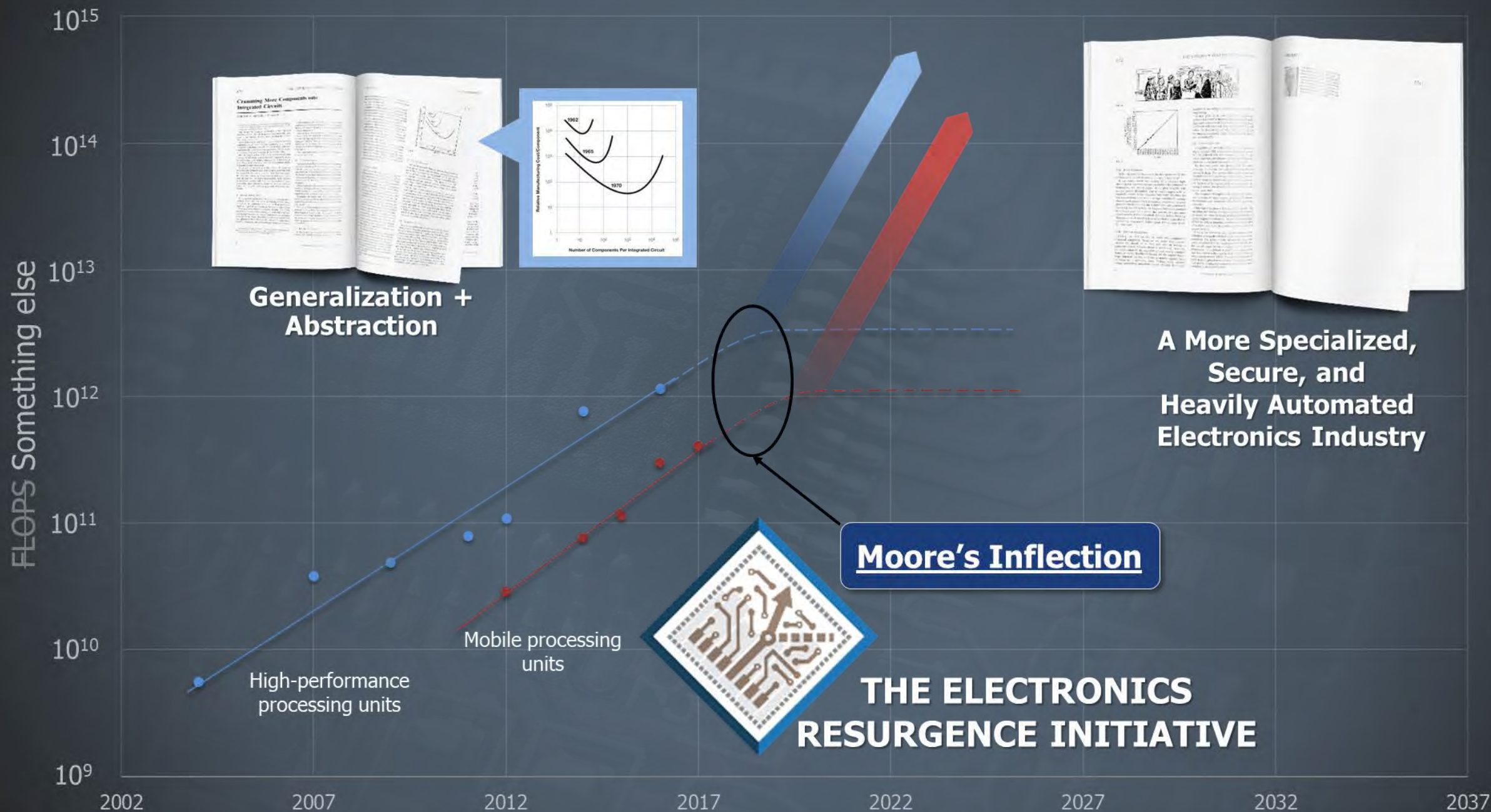
How do we lower the design barrier to specialization?



Architectures Thrust

How do we manage the complexity of specialization with new architectures?







ERI ELECTRONICS RESURGENCE INITIATIVE SUMMIT 2018 | SAN FRANCISCO, CA | JULY 23-25

CONCLUDING THOUGHTS: EVERYTHING OLD IS NEW AGAIN

- Dave Kuck, software architect for Iliac IV (circa 1975)
"What I was really frustrated about was the fact, with Iliac IV, programming the machine was very difficult and the architecture probably was not very well suited to some of the applications we were trying to run. The key idea was that I did not think we had a very good match in Iliac IV between applications and architecture."
- Achieving cost-performance in this era of DSAs will require matching the applications, languages, architecture, and reducing design cost.
- Information technology (computing to electronics) is the most important economic and security asset for any nation: Combine SW/HW/creativity to compete by running faster.

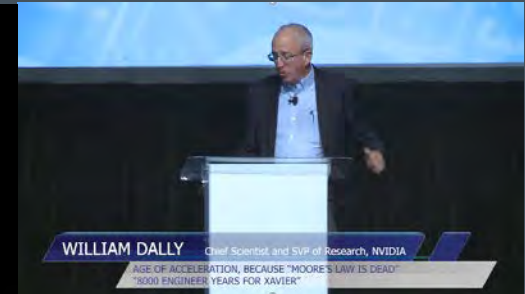


XAVIER

World's First Autonomous Machine Processor



Most Complex SOC Ever Made | 9 Billion Transistors, 250mm², 170FPK | ~8,000 Engineering Years
Diversity of Engines Accelerates Entire AI Pipeline | Designed for ADAS & AI



A USEFUL METAPHOR – OCEANS AND ISLANDS

- Ocean represents big market general purpose
- Islands represents specific purpose products
- The ocean performance level rises over time
- Their job is to be sufficiently better than the ocean

Why not use both?

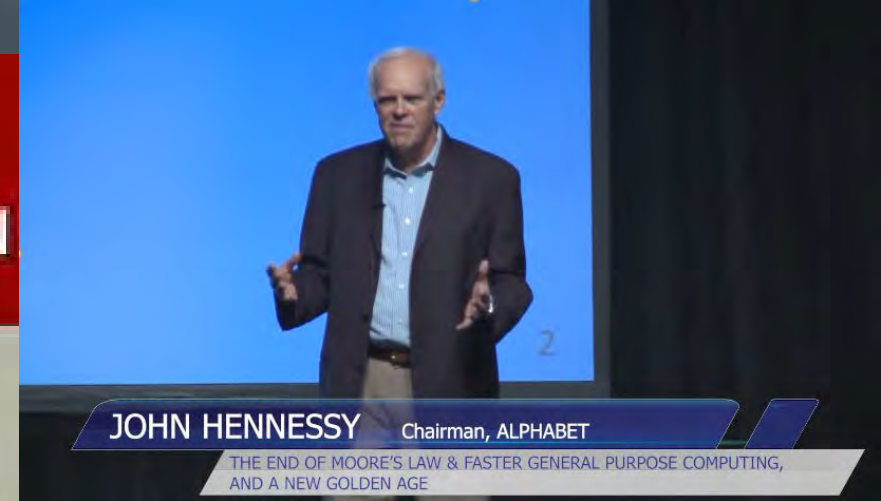


China Pulls Away in Last Four Quarters Worldwide Fabless Company Venture Capital Rounds (1-3)



CONCLUDING THOUGHTS: EVERYTHING OLD IS NEW AGAIN

- Dave Kuck, software architect for Illiac IV (circa 1975)
"What I was really frustrated about was the fact, with Illiac IV, programming the machine was very difficult and the architecture probably was not very well suited to some of the applications we were trying to run. The key idea was that I did not think we had a very good match in Illiac IV between applications and architecture."
- Achieving cost-performance in this era of DSAs will require matching the applications, languages, architecture, and reducing design cost.
- Information technology (computing to electronics) is the most important economic and security asset for any nation: Combine SW/HW/creativity to compete by running faster.

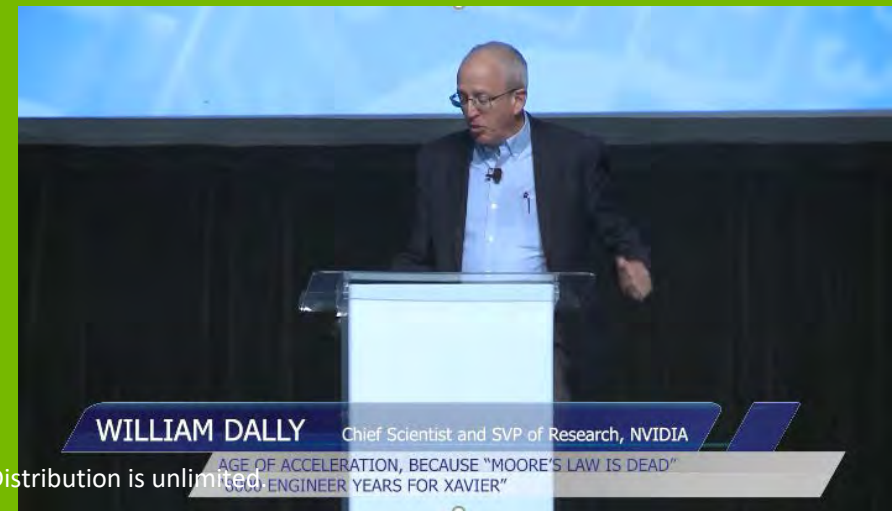


JOHN HENNESSY Chairman, ALPHABET

THE END OF MOORE'S LAW & FASTER GENERAL PURPOSE COMPUTING,
AND A NEW GOLDEN AGE

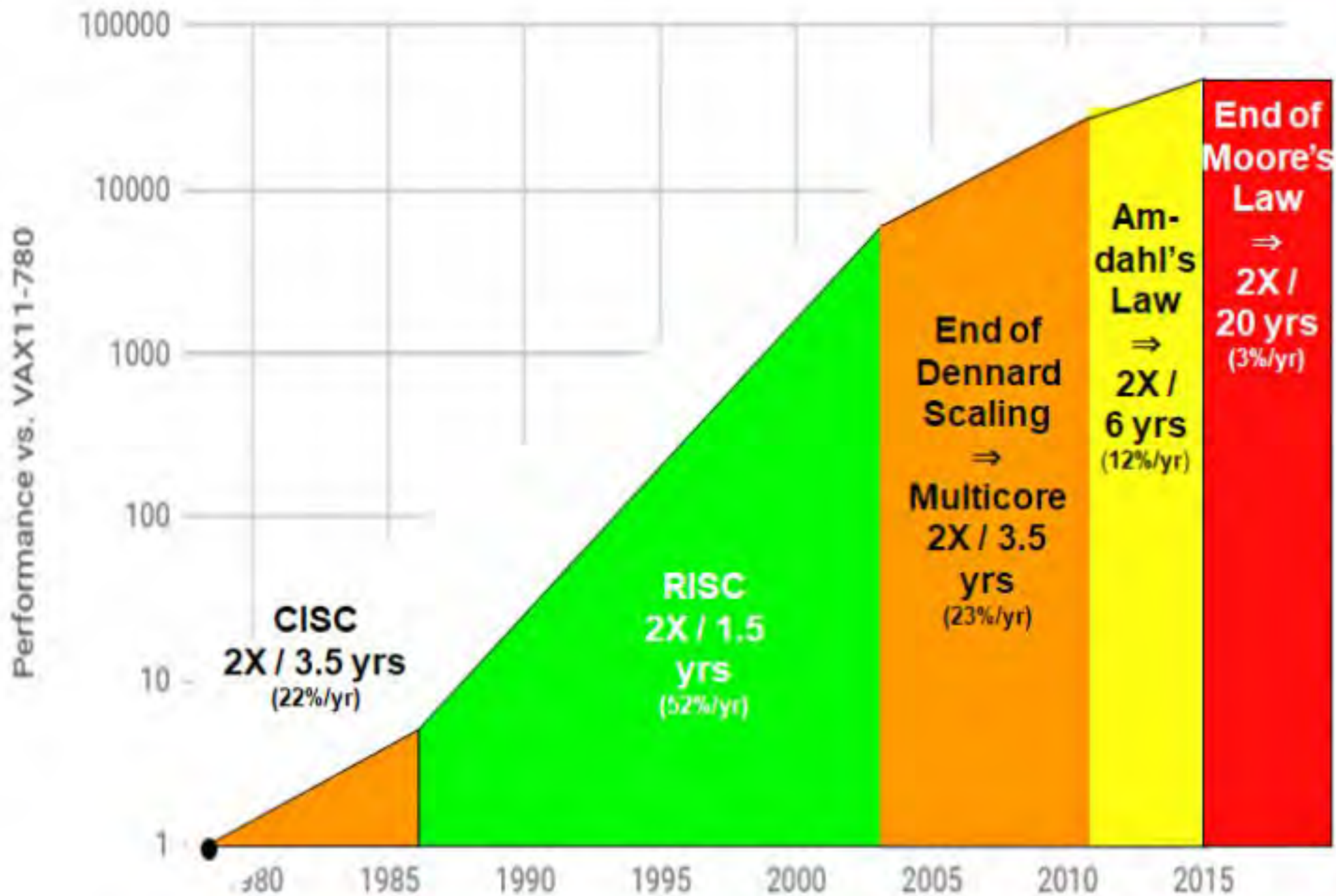
But Process Technology isn't Helping us
Anymore

Moore's Law is Dead



WILLIAM DALLY Chief Scientist and SVP of Research, NVIDIA

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited. AGE OF ACCELERATION, BECAUSE "MOORE'S LAW IS DEAD"



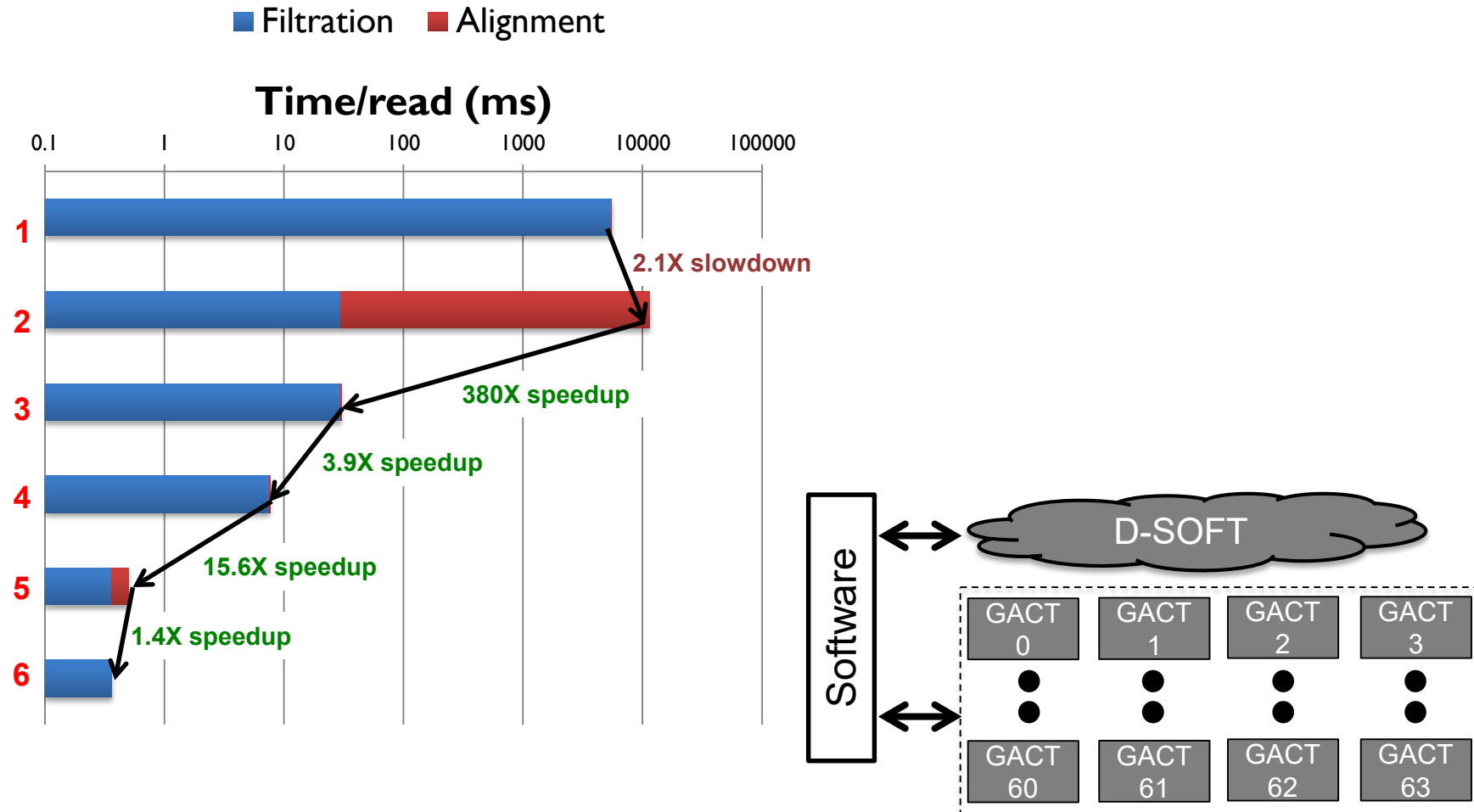
DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

Accelerators can continue scaling
perf and perf/W

Algorithm-Hardware Co-Design for Darwin

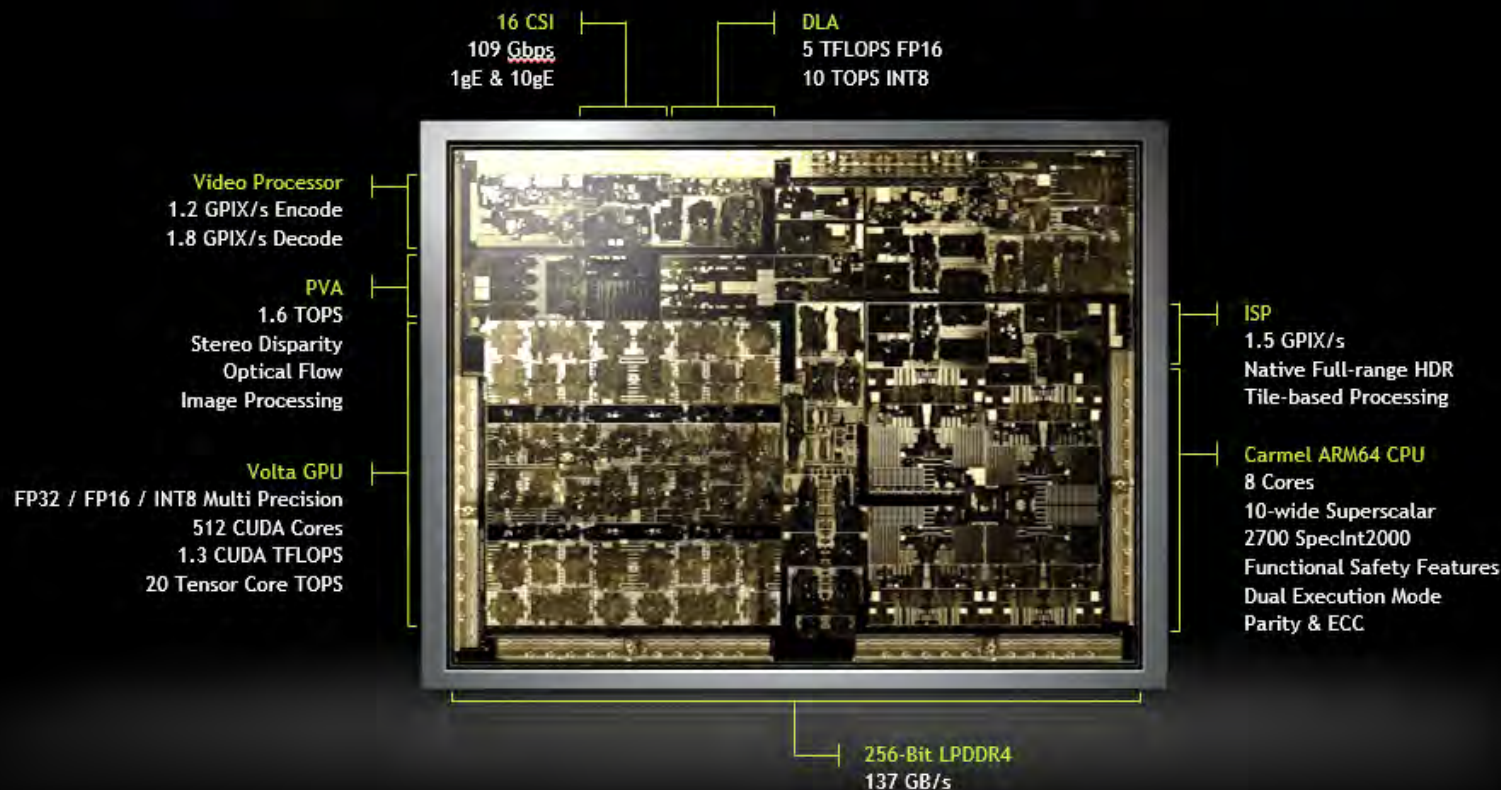
Pipeline D-Soft and GACT – now completely D-Soft limited – 1.4x

Overall 15,000x



XAVIER

World's First Autonomous Machine Processor



Most Complex SOC Ever Made | 9 Billion Transistors, 350mm², 12nFFN | ~8,000 Engineering Years
Diversity of Engines Accelerate Entire AV Pipeline | Designed for ASIL-D AV

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

A USEFUL METAPHOR – OCEANS AND ISLANDS

Ocean represents big market
general purpose

Islands represents
specific purpose products

The ocean performance level
rises over time

Their job is to be sufficiently
better than the ocean

Why not use both?



MIKE MAYBERRY

CTO & Managing Director, Intel LABS

THE FUTURE IS SO BRIGHT; THE ERA OF HETEROGENEOUS COMPUTE
"OCEANS AND ISLANDS"

00:00:03

China Pulls Away in Last Four Quarters

Worldwide Fabless Company Venture Capital Rounds (1-3)



WALDEN RHINES President & CEO, Mentor, a Siemens Business

IC GROWTH OF 22% AFTER COMPOUND ANNUAL GROWTH RATE 3%
"HAPPY DAYS ARE HERE AGAIN"

COMPLEXITY BIND

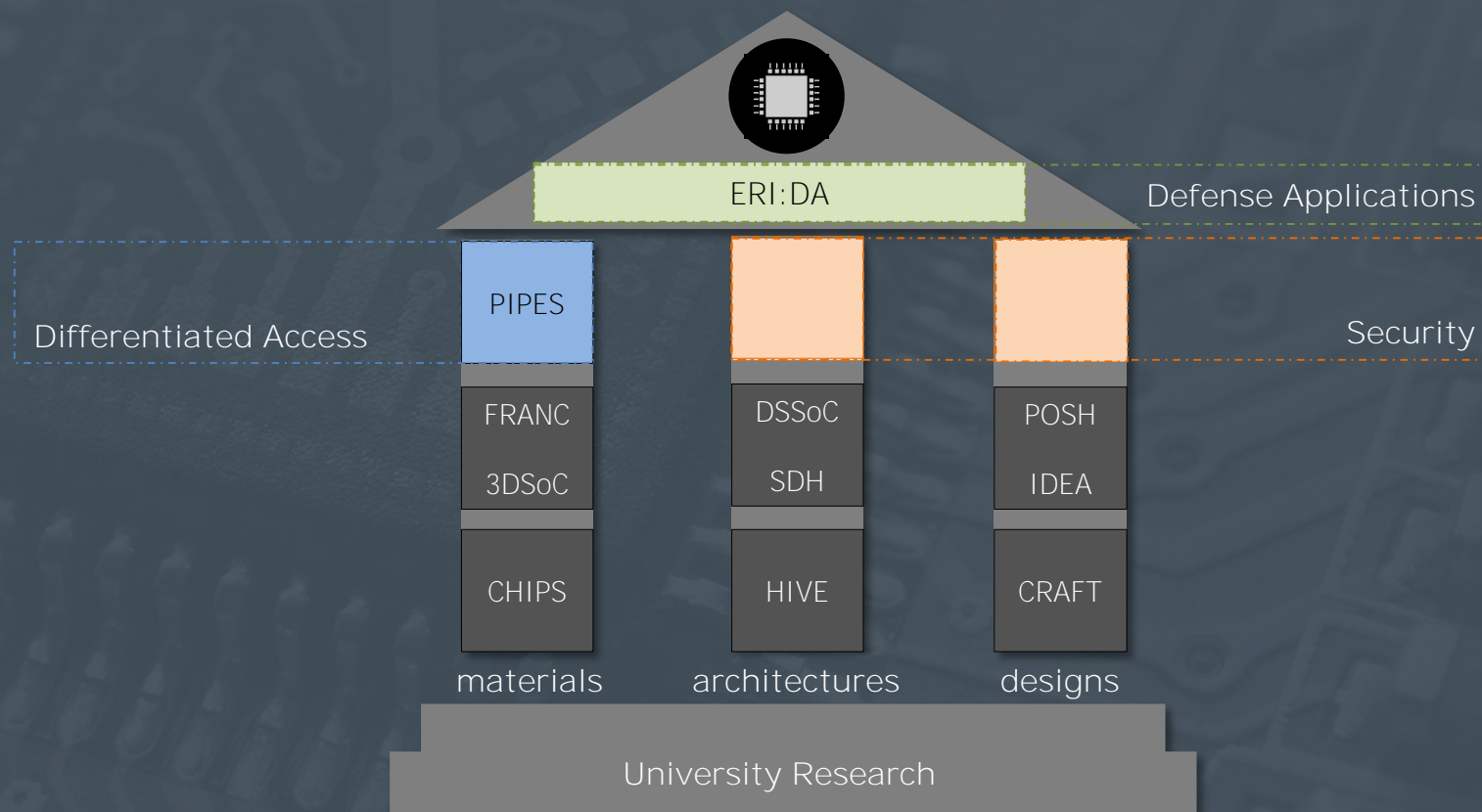


The cost of integrated circuit fabrication, design, and verification is skyrocketing, limiting innovation

The continued move towards generalization and abstraction is stifling potential gains in hardware

Digital influence is so pervasive in our society that we can't afford to have flaws in the digital foundation

Phase II







THE ELECTRONICS RESURGENCE INITIATIVE

ENABLING COLLABORATION

Richard-Duane Chambers
ERI Special Assistant
Contact: ERI_Page3@darpa.mil



ERI **ELECTRONICS
RESURGENCE INITIATIVE**
SUMMIT
2018 | SAN FRANCISCO, CA | JULY 23-25

Universities

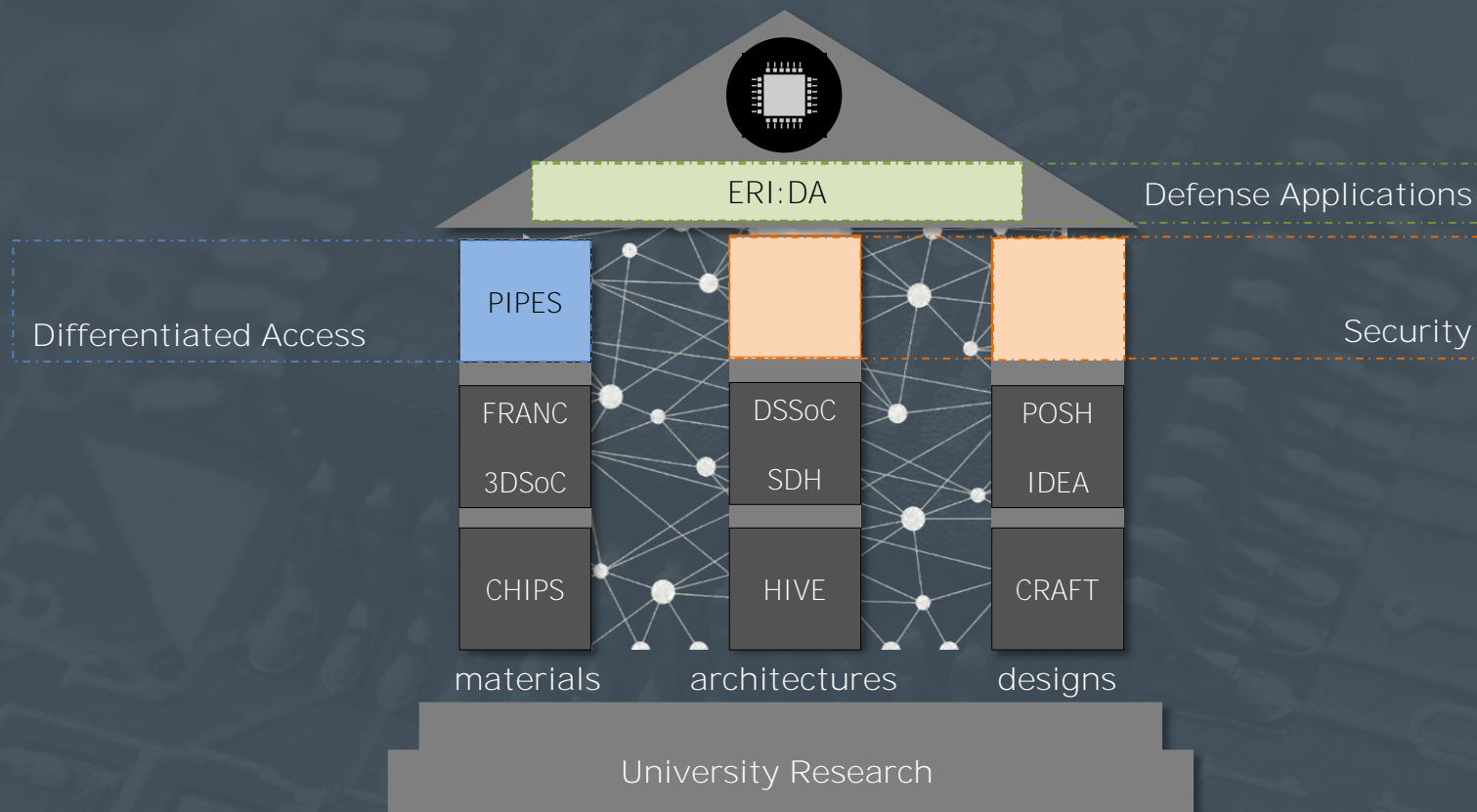
Arizona State University
Brown University
Cornell University
Georgia Tech
MIT
Princeton University
Purdue University
Stanford University
University of California
University of Illinois - UC
University of Michigan
University of Minnesota
University of Southern CA
University of Texas
University of Utah
University of Washington
Yale University

Commercial

Applied Materials
ARM
Cadence
Ferric Semiconductor
Global Foundries
IBM
Intel
Le Wiz
Micron
NVIDIA
Qualcomm
Samsung
Skywater
Synopsys
Systems & Technology
Research
TSMC
Xilinx

Defense

Army Research Lab
Boeing
General Dynamics
General Electric
HRL Laboratories
Lockheed Martin
NIST
Northrop Grumman
Oak Ridge National Lab
Raytheon
Sandia National Labs



Page 3

- 3DSOC
- FRANC
- SDH
- DSSoC
- IDEA
- POSH

Phase II

- PIPES

Foundational

- JUMP
- CRAFT
- CHIPS
- HIVE
- L2M
- MIDAS
- N-ZERO
- SSITH

Notional BAA Process

- Purpose: To accelerate the delivery of ERI -derived innovations to national security needs by demonstrating and applying emerging ERI technologies
- \$25 million total for ERI: Defense Applications additions to existing programs
- 6.3 Funding
- Pre-BAA Proposers Day (19 December 2018)
- Tentative BAA Release (15 January 2019)
- Tentative abstract deadline (5 February 2019) *
- Tentative 1st round proposals deadline (28 March 2019)
- Tentative 1st round Government selections (May 2019)
- Tentative contract awards (August 2019)

*Abstracts are highly recommended.

New opportunities to participate

Track 1 – Immediate Technology Development

- Proposal from a demonstrated national security partner
- Verifies an existing or planned relationship with an existing ERI performer
- Identifies an existing or planned technology under development in an ERI program
- Identifies a specific, revolutionary defense application that would benefit from ERI-developed technology
- Provides a detailed proposal for the technology development, demonstration, and application

Track 2 – Partnering and Technology Development

- Proposal from a demonstrated national security partner
- Provides a plan to create a relationship with an existing ERI performer within 12 months
- Identifies an existing or planned technology under development in an ERI program
- Identifies a specific, revolutionary defense application that would benefit from ERI-developed technology
- Provides a detailed proposal for the technology development, demonstration, and application

Page 3

- 3DSOC
- FRANC
- SDH
- DSSoC
- IDEA
- POSH

Phase II

- PIPES

Foundational

- JUMP
- CRAFT
- CHIPS
- HIVE
- L2M
- MIDAS
- N-ZERO
- SSITH

New opportunities to participate

Track 1 – Immediate Technology Development

- Proposal from a demonstrated national security partner
- Verifies an existing or planned relationship with an existing ERI performer
- Identifies an existing or planned technology under development in an ERI program
- Identifies a specific, revolutionary defense application that would benefit from ERI-developed technology
- Provides a detailed proposal for the technology development, demonstration, and application

Track 2 – Partnering and Technology Development

- Proposal from a demonstrated national security partner
- Provides a plan to create a relationship with an existing ERI performer within 12 months
- Identifies an existing or planned technology under development in an ERI program
- Identifies a specific, revolutionary defense application that would benefit from ERI-developed technology
- Provides a detailed proposal for the technology development, demonstration, and application

Page 3

- 3DSOC
- FRANC
- SDH
- DSSoC
- IDEA
- POSH

Phase II

- PIPES

Foundational

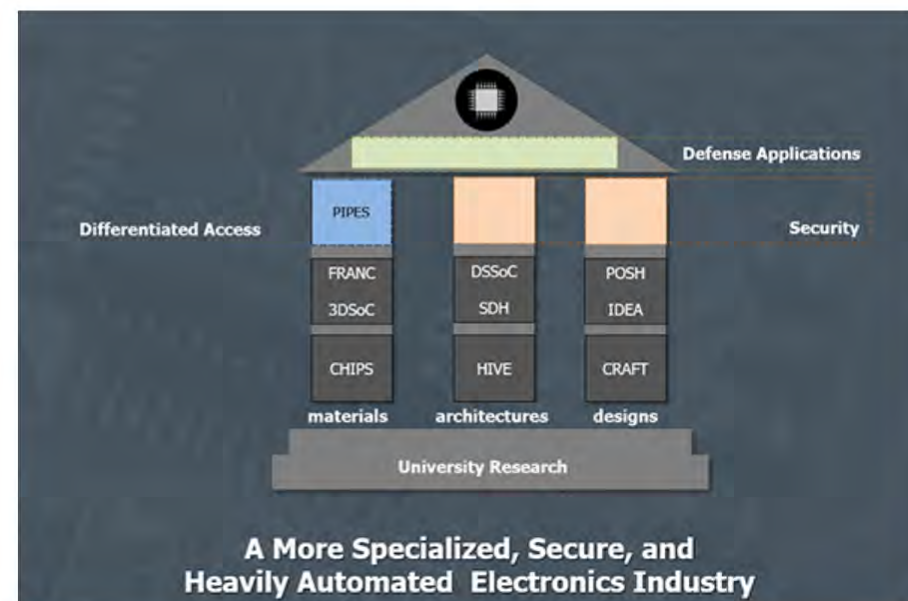
- JUMP
- CHIPS
- CRAFT
- HIVE
- L2M
- MIDAS
- N-ZERO
- SSITH

Enabling Collaboration – ERI Website

DARPA Electronics Resurgence Initiative

On June 1, 2017, the DARPA Microsystems Technology Office (MTO) [announced](#) a new Electronics Resurgence Initiative (ERI) to ensure far-reaching improvements in electronics performance well beyond the limits of traditional scaling. ERI draws on new and existing DARPA programs to make a significant investment into enabling circuit specialization and managing complexity. Building on the tradition of other successful government-industry partnerships, ERI aims to forge forward-looking collaborations among the commercial electronics community, defense industrial base, university researchers, and the DoD to create a more specialized, secure, and heavily automated electronics industry that serves the needs of both the domestic commercial and defense sectors.

ERI Overview and Structure



OPEN OPPORTUNITIES

[HR001119S0004: Photonics in the Package for Extreme Scalability \(PIPES\)](#)

ONGOING ERI PROGRAMS

Phase II

[Photonics in the Package for Extreme Scalability \(PIPES\)](#)

Page 3

[Three Dimensional Monolithic System-on-Chip \(3DSOC\)](#)

[Foundations Required for Novel Compute \(FRANC\)](#)

[Software Defined Hardware \(SDH\)](#)

[Domain-Specific System on Chip \(DSSoC\)](#)

[Intelligent Design of Electronic Assets \(IDEA\)](#)

[Posh Open Source Hardware \(POSH\)](#)

Foundational

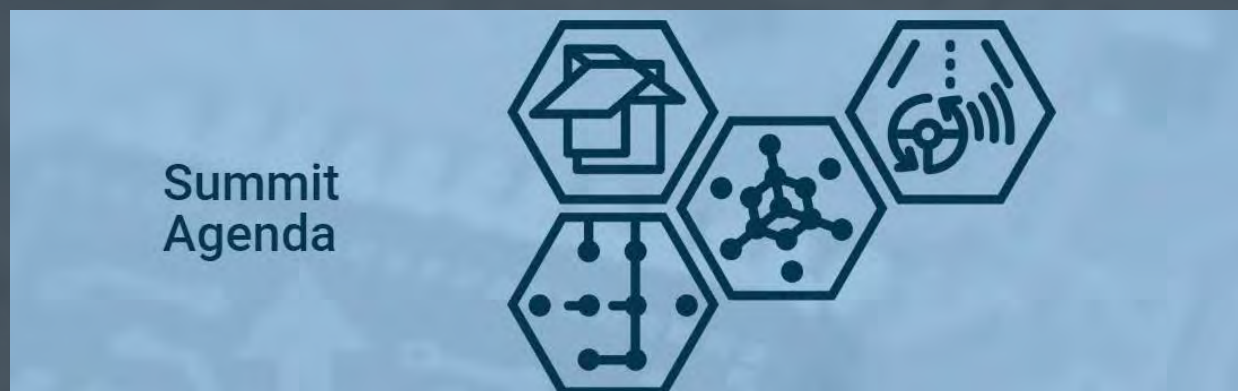
[Circuit Realization At Faster Timescales \(CRAFT\)](#)

[Compact Heterogeneous Integration and IP Reuse Strategies \(CHIPS\)](#)

[Hierarchical Identify Verify Exploit \(HIVE\)](#)

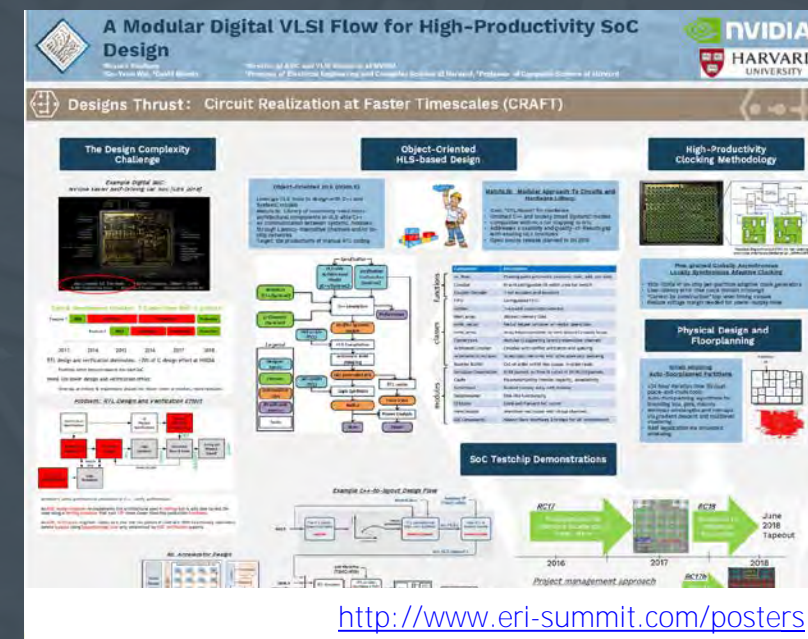
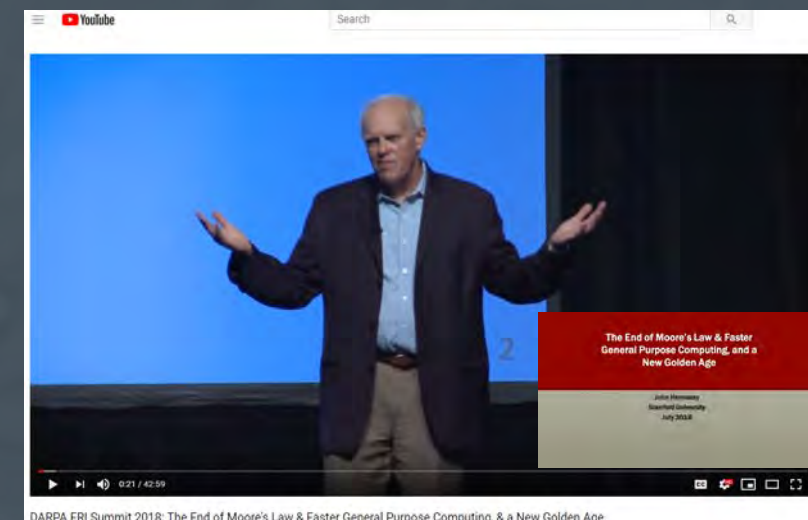
<https://www.darpa.mil/work-with-us/electronics-resurgence-initiative>

Enabling Collaboration – ERI Summit Materials



2:30 PM	<p><u>Welcome and Announcement of ERI "Page 3" Teams</u> Dr. Jay Lewis, Deputy Director, Microsystems Technology Office, DARPA</p>
3:15 PM	<p><u>Introduction of Joint University Microelectronics Program (JUMP) Focus Areas</u> Dr. Linton Salmon, Program Manager, DARPA MTO Dr. Anthony Rowe, Associate Professor, Electrical & Computer Engineering, Carnegie Mellon University Dr. Valeria Bertacco, Professor, Electrical Engineering, University of Michigan Dr. Suman Datta, Freimann Chair of Engineering Professor, Notre Dame University Dr. Kaushik Roy, Professor, Electrical & Computer Engineering, Purdue University Dr. Mark Rodwell, Professor, Electrical & Computer Engineering, University of California, Santa Barbara Dr. Tajana Rosing, Professor, Computer Science and Engineering, University of California, San Diego</p>
4:30 PM	<p><u>Science and Policy at the End of Moore's Law</u> Dr. Erica Fuchs, Professor, Engineering and Public Policy, Carnegie Mellon University</p>
5:00 PM	<p><u>Impact of Commercial Partnership with the DoD</u> Mr. Tom Beckley, Senior Vice President & GM of Custom IC & PCB Group, Cadence Design Systems</p>
5:30 PM	<p><u>Reception and Networking</u></p>
7:00 PM	<p><u>Adjourn</u></p>

<http://www.eri-summit.com/agenda>



<http://www.eri-summit.com/posters>

Enabling Collaboration – ERI:DA Proposers Day

Start	Session	
7:45	<i>Check-In (Online registration required before deadline. No onsite registration.)</i>	
9:00	Greg Woosley	Security Discussion
9:05	Bill Chappell	DARPA ERI and Defense Applications
9:20	Richard-Duane Chambers	Enabling Collaboration
9:35	Michael Blackstone	Contracting Guidance
10:05	Wade Shen	SDH
10:30	Thomas Rondeau	DSSoC
10:55	Collaboration Period / Poster Session	
11:30	Lunch	
12:30	Andreas Olofsson	CRAFT, IDEA, POSH, CHIPS
13:10	Young-Kai Chen	FRANC
13:35	Jay Lewis	3DSoC, JUMP
14:00	Collaboration Period / Poster Session	
14:30	Q&A Period (<i>Notecard submission</i>)	
15:30	Collaboration Period	
16:00	END	

- Collaboration Periods
- Open Rooms for Discussions
- Posters
- Q&A
- Mailing List

Contact: ERI_Page3@darpa.mil





**ERI:DA
HR001119S0018
(ERI Phase II)**

**Proposers Day
(Pre-BAA)**

December 19, 2018

**Michael Blackstone
Contracting Officer
DARPA Contracts Management Office**





Proposers Day Disclaimer

- **Plenty of good information is made available to potential proposers to help clarify program goals/objectives and proposal preparation instructions - those things that are (or may be) stipulated in the BAA**

However:

- Only the information/instructions in the BAA counts
- Proposals will only be evaluated in accordance with the instructions provided in the BAA
- Any response provided by the Government in the FAQ that's different than what is provided in the BAA will be made formal by an amendment to the BAA
 - Such responses will make note of an impending BAA amendment
- **The ERI:DA BAA has not been finalized or published so things could change. So be sure to read the BAA once published**
- **Being a pre-BAA event allows potential proposers to speak freely with DARPA PMs and for such input to potentially have a bearing on the BAA**
 - ✓ Pre-BAA questions: ERI_page3@darpa.mil



BAA Overview

- **BAA allows for a variety of technical solutions and award instrument types**

- The BAA defines the problem set, the proposer defines the solution (and SOW)
- Allows for multiple award instrument types:

- Procurement Contract
- Other Transaction (OT) Agreement
(No Grants or Cooperative Agreements)

**6.3 Funding Only (Adv. Tech. Dev.)
Restricted Research (all tiers)**

- **DARPA Scientific Review Process**

- Proposals are evaluated on individual merit and relevance as it relates to the stated research goals/objectives rather than against one another (there is no common statement of work)
- Selections will be made to proposers whose proposals are determined to be most advantageous to the Government, all factors considered, including potential contributions to research program and availability of funding

- **Government may select for negotiation all, some, or none of the proposals received**
- **Government may accept proposals in their entirety or select only portions thereof**
- **Government may elect to establish portions of proposal as options**



BAA Process/Timeline

(Notional)

1. **Pre-BAA Proposers Day is conducted (19 December 2018)**
 2. **BAA is released (~ 15 January 2019)**
 3. **Abstracts are due (~ 5 February 2019)**
 4. **Government abstract responses (~ 26 February 2019)**
 5. **Proposals are due/submitted (~ 28 March 2019)**
 - **Proposals can be submitted beyond for up to ~180 days to a year**
 6. **Proposals are reviewed for BAA compliance**
 - Noncompliant proposals are not reviewed (and cannot be selected)
 7. **Government conducts Scientific Review Process**
 - Clarification requests may be sent to various proposers
 8. **Government sends out select/non-select letters (~ 12 May 2019)**
 - All proposers who submit a compliant proposal may request an Informal Feedback Session
 9. **Contracting Officer initiates negotiations (awards by August 2019)**
- Timeline Annotations:**
- A bracket between steps 2 and 3 indicates a duration of **21 days**.
 - A bracket between steps 4 and 5 indicates a duration of **21 days**.
 - A bracket between steps 5 and 6 indicates a duration of **30 days**.
 - A bracket between steps 7 and 8 indicates a duration of **45 days**.
 - A bracket between steps 8 and 9 indicates a duration of **90 days**.



Eligibility Issues

- **All interested/qualified sources** may respond subject to the parameters outlined in BAA (**such as, for ERI:DA, accepting use of 6.3 funding**)
- Foreign participants/resources may participate to the extent allowed by applicable Security Regulations, Export Control Laws, Non-Disclosure Agreements, etc. (No classified proposals anticipated)
- FFRDCs and Government entities:
 - Are not prohibited by the BAA from proposing
 - Are, however, subject to applicable direct competition limitations
 - Are, however, required to demonstrate eligibility (sponsor letter)
 - **The burden to prove eligibility for all such team members rests with the proposer**
 - **All elements of a proposal (tech and cost, prime and subs – even FFRDC team members) must be included in the prime's submission**
- Real and/or Perceived Conflicts of Interest:
 - Identify any conflict/s
 - If any are identified, a mitigation plan must be included



Teaming & Contracting Considerations

● Teaming Alternatives

- Industry Partner (ERI:DA) Prime / ERI Performer Subcontractor Relationship
 - Preferred when instrument type for both organizations align
 - Preferred when funding type – 6.3/restricted – at both tiers align

- Industry Partner (ERI:DA) Prime / ERI Performer Associate Contractor Relationship
 - Required when instrument type at both tiers do not align
 - Required when funding type – 6.3/restricted – at both tiers do not align

Alignment = Industry Partner (ERI:DA) and ERI Performer can work under a procurement contract or OT with 6.3 funding (restricted/non-fundamental) as prime and subcontractor

Nonalignment = Industry Partner (ERI:DA) requires a procurement contract and the ERI Performer requires an OT, and/or

Industry Partner (ERI:DA) can accept use of 6.3 (restricted/non-fundamental) funding but the ERI Performer can only accept 6.2 (fundamental) funding



Teaming & Contracting Considerations

- **Proposal must align with the Track and Teaming Approach:**

- Track 1 or Track 2?
- ERI:DA Prime / ERI Performer Sub Relationship?
- ERI:DA Prime / ERI Performer Associate Contractor Relationship?
-
- Track 1 proposals must include a complete technical approach and cost proposal?
 - This applies no matter the teaming approach used (prime/sub or ACA)
-
- Track 2 proposers must include (only) the technical approach and cost proposal as it relates to the prime contractor's task activity (because the ERI Performer will be unknown at this time)
- Track 2 performers will be required to define and establish the teaming relationship during Phase 1 (prior to Phase 2 turn on)

This will result in one of the two following approaches to pull the ERI Performer into Phase 2:

1. Modification to the ERI:DA award instrument to add subcontract tasks & costs
2. Modification to the ERI Performer award instrument to add tasks & costs

Reminder: Track 1 – Existing/Planned relationship with an existing ERI performer



Teaming & Contracting Considerations

● Associate Contractor Relationship

- ERI:DA Industry Partner and ERI Performer have no contractual prime/sub relationship
- Each operates under a separate (Prime) contract with the Government
- ERI:DA Industry Partner and ERI Performer establish a collaboration relationship via an Associate Contractor Agreement (ACA):
 - ERI Performer's DA effort/tasks would be added to their existing ERI Phase 1/Page 3 instrument using the appropriate funding type (such as 6.2 if a university)
 - The ACA sets the basic collaboration relationship ground-rules to ensure both parties agree to work together to meet the defined project goals and objectives (share data)
 - Track 1 – ACA must be established/signed prior to award of the ERI:DA Industry Partner contract award and ERI Performer task-add modification award (Preaward)
 - Track 2 – ACA must be established/signed during Phase 1, prior to ERI: DA Industry Partner Phase 2 option exercise and ERI Performer task-add modification award (Postaward)
- The Government is not a party to the ACA (does not sign it – only the performers sign it)
 - Contracting Officer gets a copy for the file as verification purposes only



Teaming & Contracting Considerations

		TA1 and TA2 Available Contracting Options <i>between:</i>	
Desired contract type <i>(DARPA & Industry Partner)</i>	Existing contract type <i>(DARPA & Existing ERI Performer)</i>	Industry Partner & ERI Performer	DARPA & Existing ERI Performer
FAR-based contract	FAR-based contract	Prime-Sub ACA	Contract modification <i>(as needed)</i>
	OT	ACA	
OT	FAR-based contract	ACA	
	OT	Prime-Sub ACA	

- ✓ In the BAA, keep a very close eye on proposal preparation instructions - Track 1 and Track 2 ERI:DA (Industry Partners) proposers will have different requirements regarding technical scope/tasks and pricing associated with existing ERI Performer team members as summarized in previous slide



Proposal Abstracts

- **Abstracts are highly encouraged:**
 1. They minimize unnecessary effort in proposal preparation and review
 2. They reduce the potential expense of preparing an out of scope proposal

- **The abstract provides a synopsis of the proposed project (tech and budget)**

- **Government will reply by letter with one of two possible responses:**
 1. Encourage full proposal, and may provide feedback
 2. Discourage full proposal, and will provide rationale (may provide feedback)
 - DARPA will not communicate further (verbally or in writing)

- **Regardless of DARPA's response to an abstract, proposers may submit a full proposal**
 - DARPA will review all full proposals submitted without regard to abstract recommendation/feedback



Full Proposal Preparation

● E. National Security Impact Statement

➤ This proposal topic is relatively new to MTO BAAs

- How the proposed work contributes to U.S. national security and U.S. technological capabilities. The proposer may also summarize previous work that contributed to U.S. national security and U.S. technological capabilities.
- Plans and capabilities to transition technologies developed under this effort to U.S. national security applications and/or to U.S. industry. The proposer may also discuss previous technology transitions to the benefit of U.S. interests.
- Any plans to transition technologies developed under this effort to foreign governments or to companies that are foreign owned, controlled or influenced. The proposer may also discuss previous technology transition to these groups.
- How the proposer will assist its employees and agents performing work under this effort to be eligible to participate in the U.S. national security environment.



Communications

- **Prior to Receipt of Proposals (Solicitation Phase):** No restrictions, however Gov't (PM/PCO) shall not dictate solutions or transfer technology
 - Typically handled through the FAQ
- **After Receipt of Proposals/Prior to Selections (Scientific Review Phase):** Limited to Contracting Officer or BAA Coordinator (with approval) to address clarifications requested by the review team
 - Proposal cannot be changed in response to clarification requests
- **After Selection/Prior to Award (Negotiation Phase):** Negotiations are conducted by the Contracting Officer
 - PM and/or COR typically tasked with finalizing the SOW (with PI)
 - PM and/or COR typically involved in any technical discussions (i.e., partial selection discussions)
 - Pre-award costs will not be reimbursed unless a pre-award cost agreement is negotiated prior to award
- **Informal Feedback Sessions (Post Selection):** May be requested/provided once the selection(s) are made
 - If made on a timely basis (~2 wks after letter), all requests will be accepted



Pitfalls That Delay (or prevent) Proposal Review

- **Failure to submit proposal on time**
 - There is a safety net built in for this BAA (rolling submissions after the initial due date) - but it is not a guarantee as funding may be exhausted during the initial round of selections

- **Failure to submit using the correct mechanism (noncompliant!)**
 - DARPA BAA site only (Procurement Contracts & OTs Only)
 - Click “Finalize Full Proposal” button or it does not get submitted
 - Pls must keep an eye on this if somebody else in your organization is submitting

- **Failure to submit both proposal volumes (noncompliant!)**
 - Volume 1, Technical/Management
 - Volume 2, Cost

- **Pages beyond the page limitation (tech prop) – pages will not be reviewed**

- **ROM/s instead of full subcontract cost proposal/s (noncompliant!)**
 - “I didn’t have time to get the subcontract proposal/s” will not change the outcome
 - This is a competition – we won’t select what we don’t understand



Software Defined Hardware

For data intensive computation

Wade Shen
DARPA I2O
December 2018



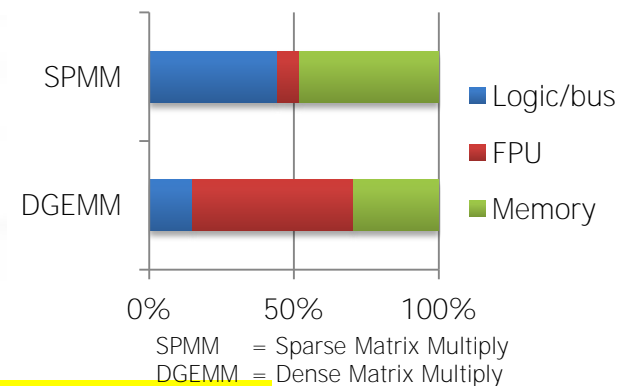
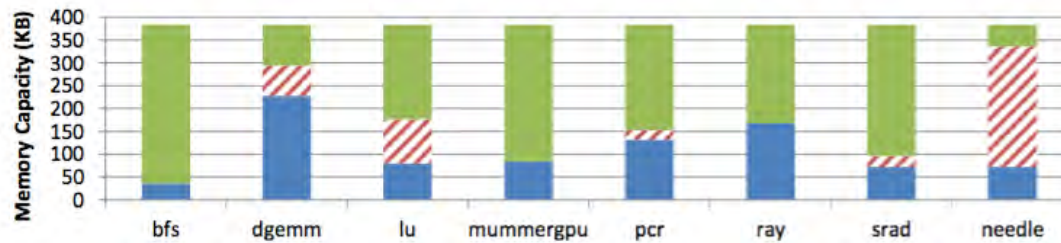
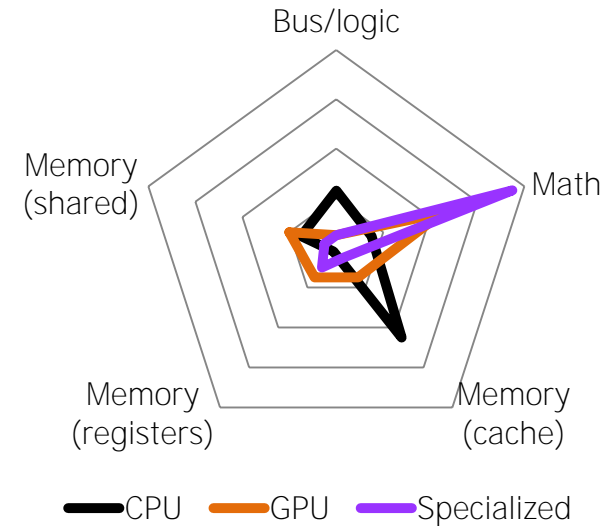
Goal Statement

Build runtime reconfigurable hardware and software that enables near ASIC performance (within 10x) without sacrificing programmability for data-intensive algorithms.



Problem

- Processor design trades
 - Math/logic resources
 - Memory (cache vs. register vs. shared)
 - Address computation
 - Data access and flow

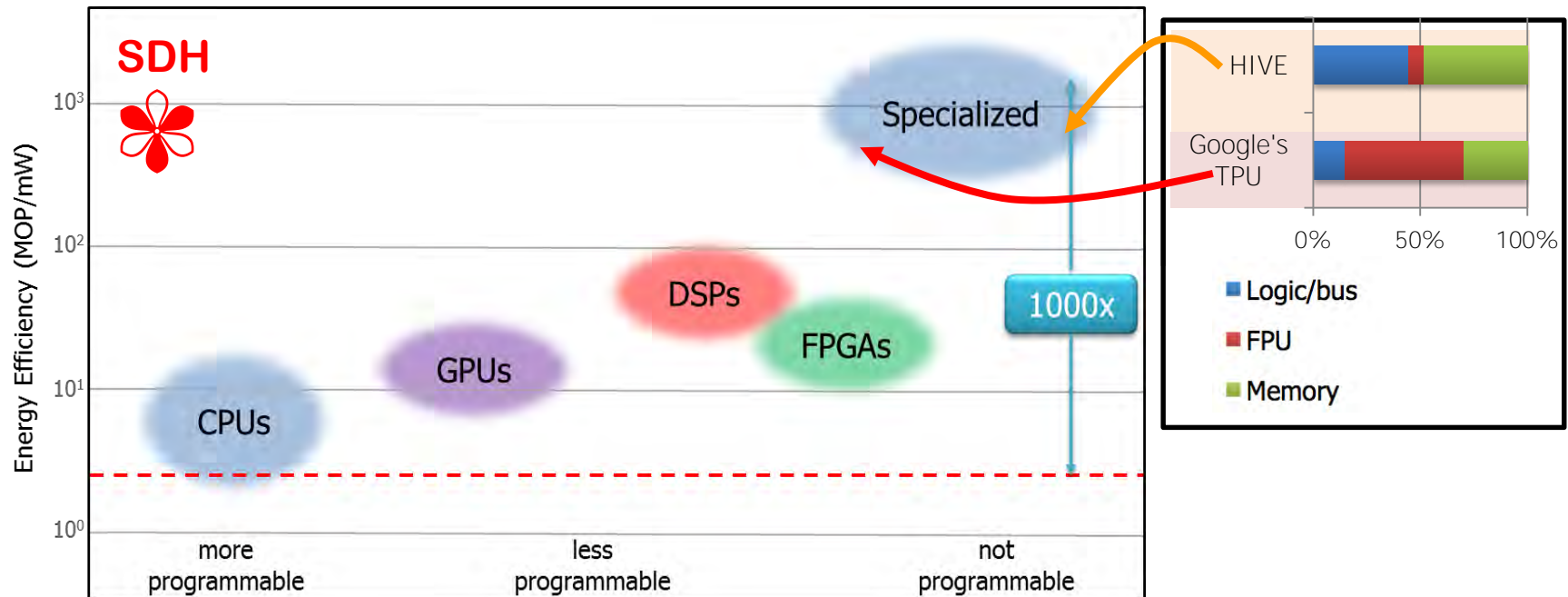


No one hardware efficiently solves all problems well



SDH: Runtime optimization of software and hardware

For data intensive computation



Today: HW design specialization

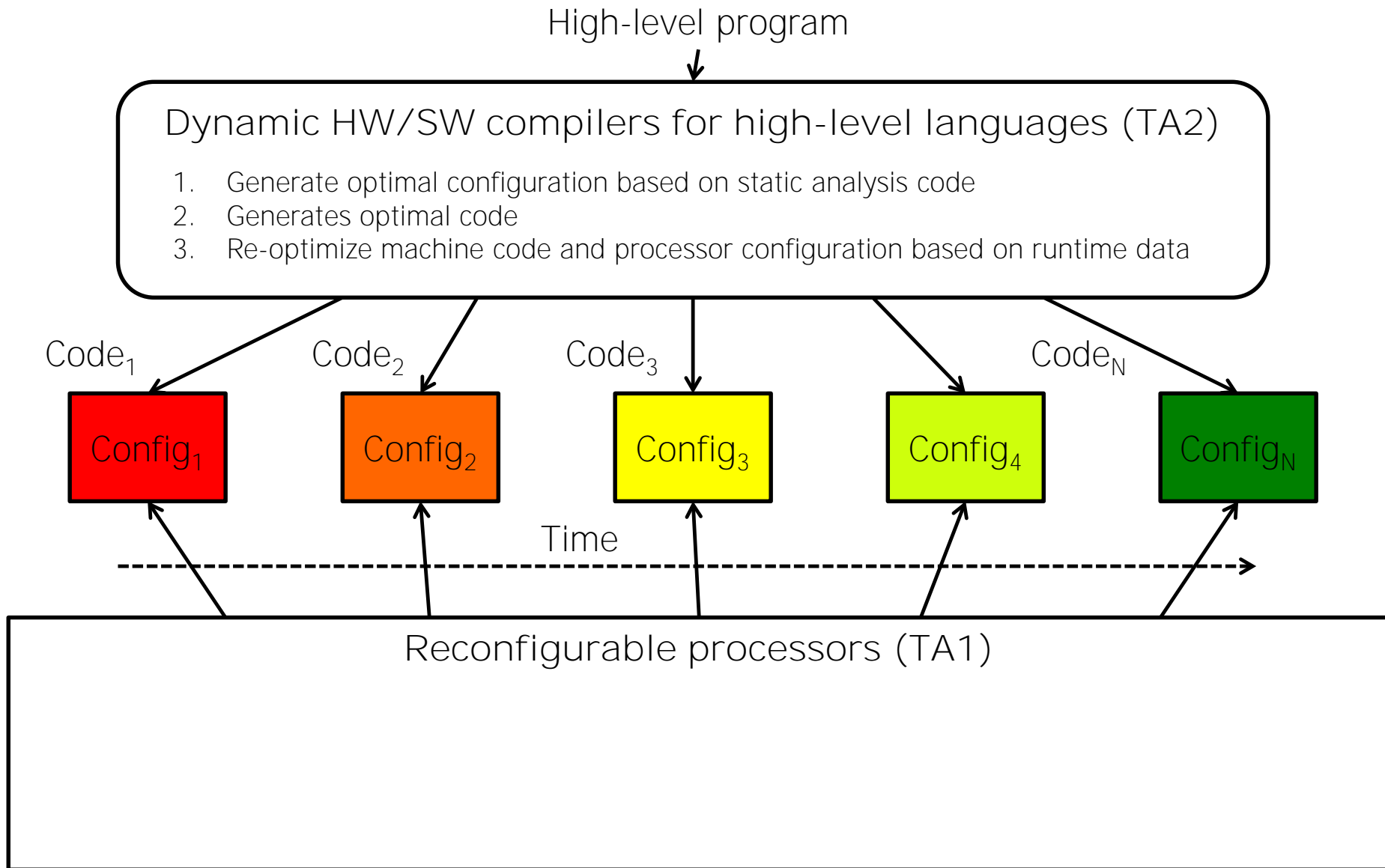
- One chip per algorithm
 - Chip design expensive
- Not reprogrammable
- Can't take advantage of data-dependent optimizations

Tomorrow: Runtime optimization of hardware and software

- One time design cost



Software-defined hardware





SDH Realization

TA2

Many Core
Hybrids

CGRA Hybrids

Just-in-time
synthesis

Offline data-
driven
optimization

Fast, low-energy
interconnect

Programmable memory
controllers

Application-programmable
memories

Code mining for kernel
optimization

Runtime type specialization
inference

ML-based processor
configuration

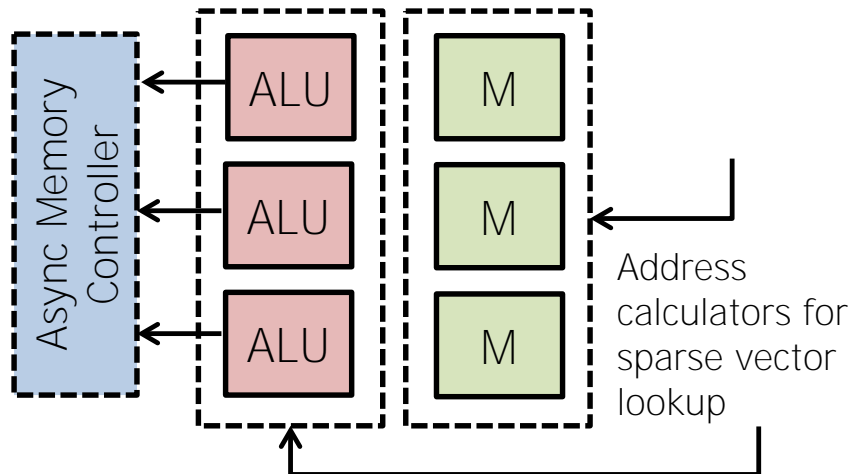
Near Memory Compute

Foundational technology



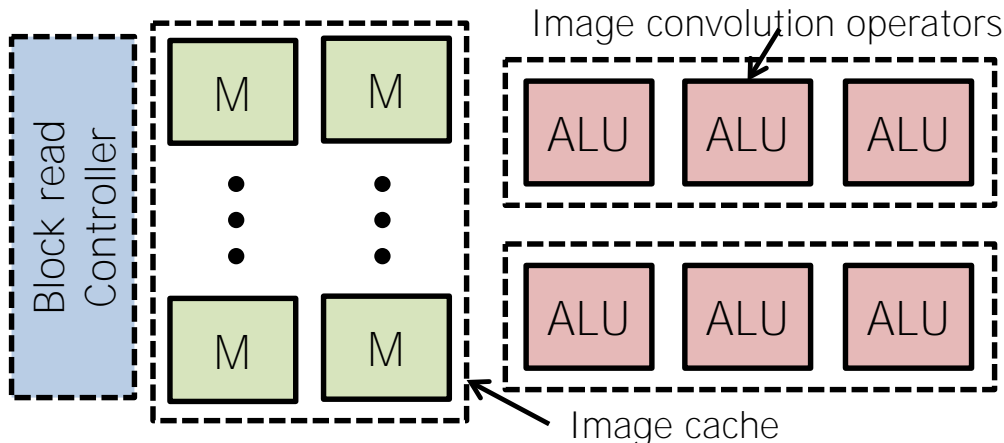
TA1: Reconfigurable processors

Graphicionado: graph search engine



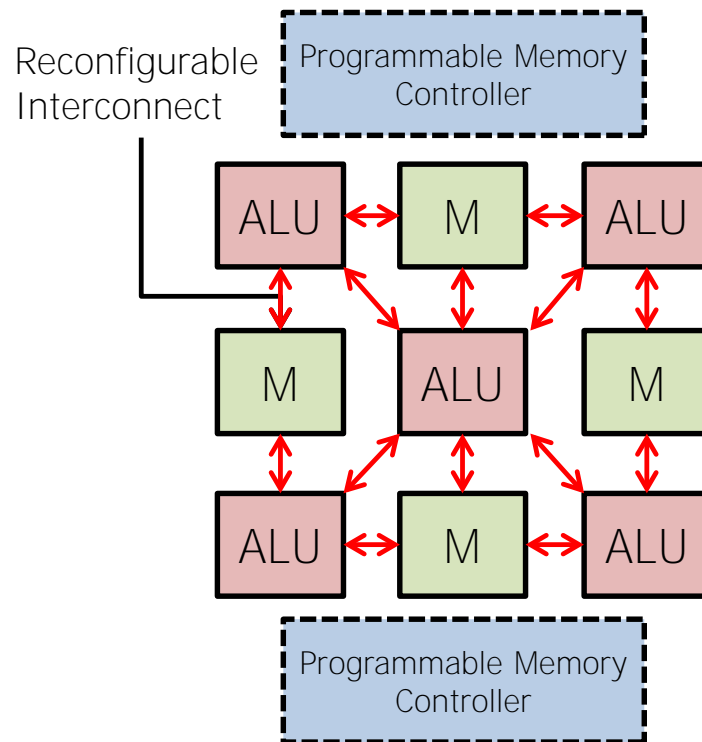
Performance: 157M edges/s/W search (BFS)

Eyeriss: Image neural net engine



Performance: 250 images/s/W (AlexNet)

SDH

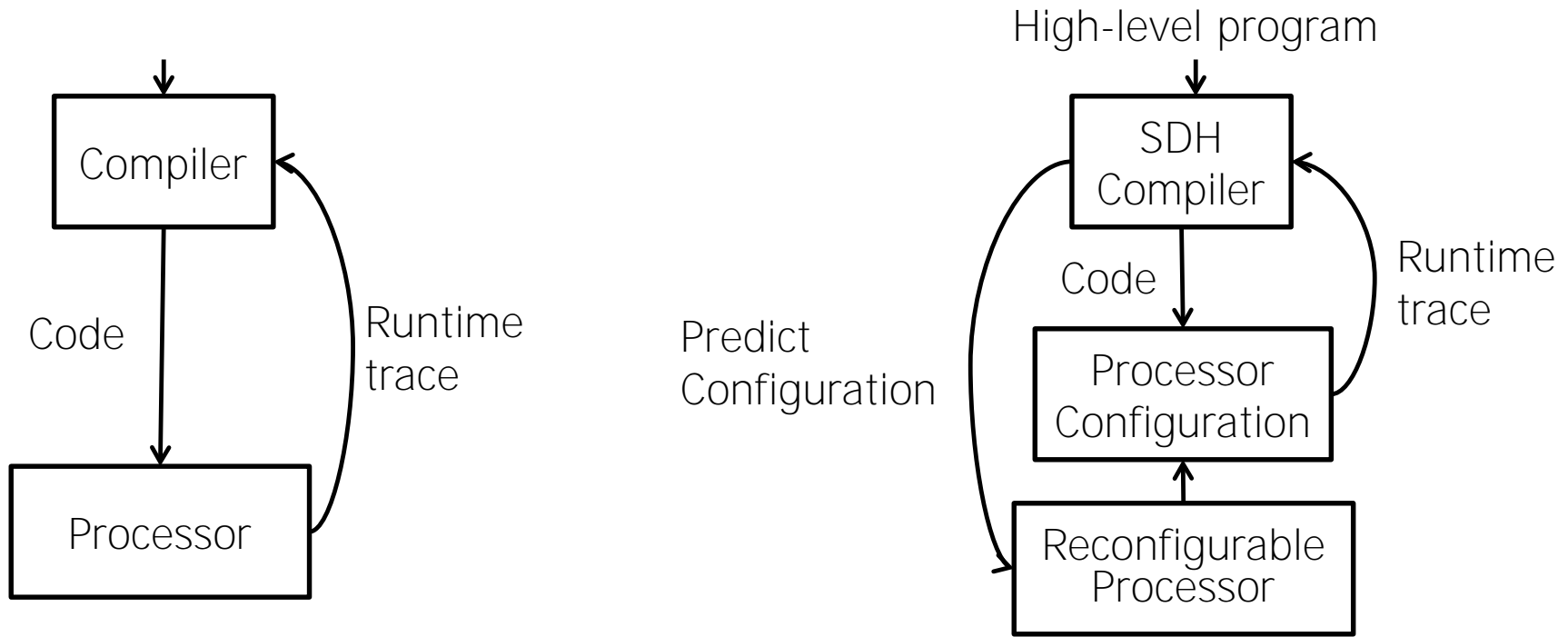


Plasticine: Stanford Seedling

- Graph search: 102M edges/s/W
- Image recognition: 130 images/s/W



TA2: Compilers to build hardware and software



- Compilers generate optimal code via static analysis + tracing methods
 - Assume static processor configuration, compile code, run, trace, recompile
- SDH compilers don't assume a static processor configuration
 - Generates optimal configuration/code given program + data
 - Problem: Resources and architecture optimization space is large
- Solution:
 1. Configure initial processor configuration, compile code, run and trace, then
 2. Predict best configuration via reinforcement learning/stochastic optimization



Problem evaluation and goals

- USG team will create a benchmark suite of machine learning, optimization, graph and numeric applications
 - Subset of 100+ programs from D3M/HIVE programs
 - Implementations for GPU and CPU
 - Subset optimized for ASIC (FPGA proxy)
- Metrics:
 - Speedup/power relative to ASIC and general purpose processors
 - Programmability:
 - Metric: time to solution for SDH vs. NumPy/Python
- Target outcomes:

Some problems from D3M/HIVE

Logistics optimization

Stochastic A/V search

Threat tracking/activity recognition

Entity resolution, link/role prediction from comm/intel graphs

Building function from satellite imagery

Crop yield prediction from satellite and weather data

Network attribution of troll amplification in social media

Multi-GMTI threat tracking and activity recognition/prediction

	vs. CPU	vs. ASIC	vs. ASIC (sparse math, graphs)	Programmability
Phase 1	100-300x	within 10x	2x	within 3x
Phase 2	500-1000x	within 5x	8-10x	~1x



SDH performers

- TA1
 - Intel
 - Qualcomm
- TA1 & TA2
 - NVIDIA
 - Princeton University
 - Stanford University
 - University of Michigan / ARM
 - University of Washington
- TA2
 - Georgia Tech Research Institute
 - Systems & Technology Research



TA1

Network of CGRA

[U. Washington](#)

CGRA Hybrids

[U. Washington](#), [Princeton](#), [Stanford](#)

Many Core Hybrids

[U. Michigan](#), [Qualcomm](#)

Fast, low-energy interconnect

[U. Michigan](#), [U. Washington](#)

Near Memory Compute

[Princeton](#), [U. Washington](#)

Programmable memory
controllers

[Stanford](#), [Intel](#), [Qualcomm](#)

Existing Small-scale Prototype

[Stanford](#)

TA2

Code mining for optimization

[STR/Purdue](#), [GA Tech](#)

Approximation synthesis

[Stanford](#), [STR/Purdue](#)

Declarative code synthesis

[Stanford](#)

Runtime type inference

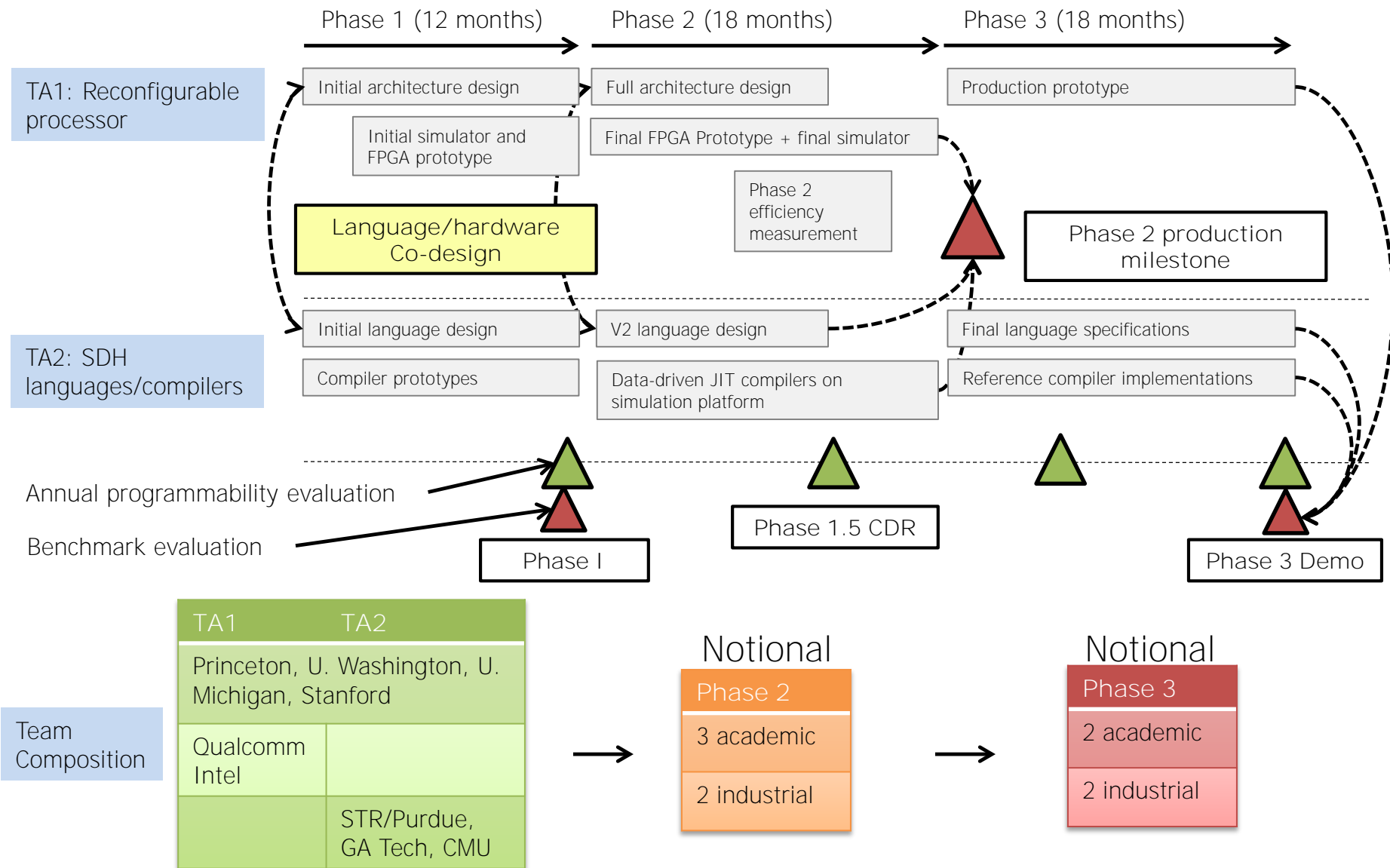
[GA Tech](#), [STR/Purdue](#)

Continuous resynthesis

[U. Washington](#), [CMU](#), [Stanford](#)

Data programming

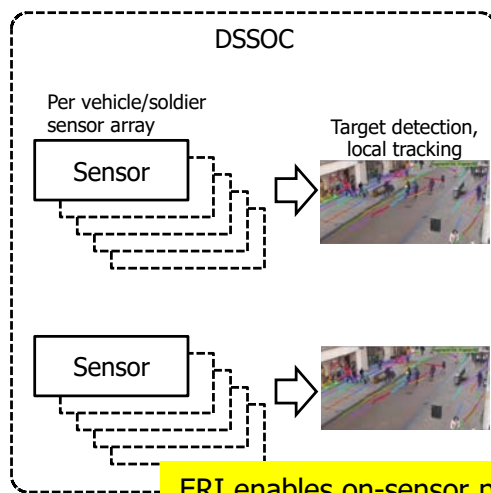
[Princeton](#)



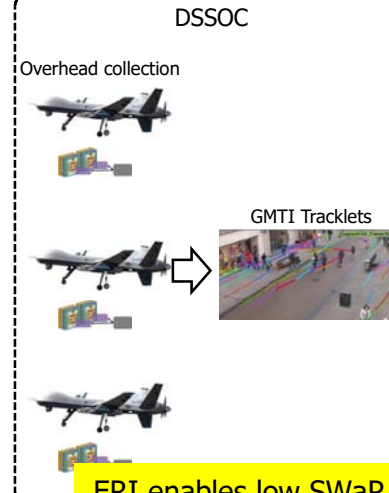
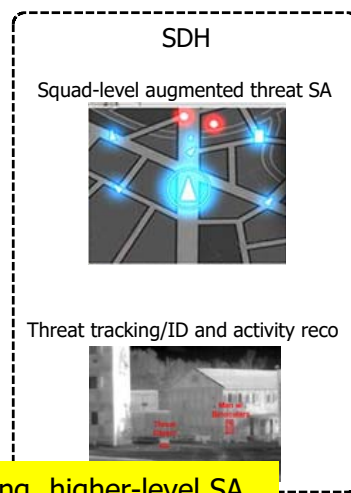


Who cares?

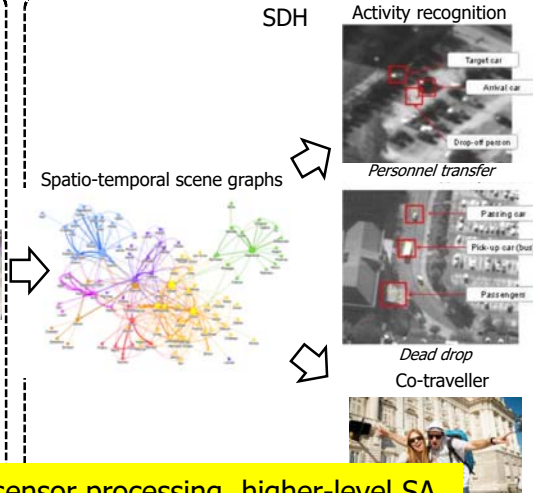
Sensors



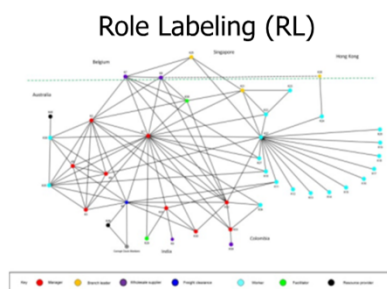
ERI enables on-sensor processing, higher-level SA



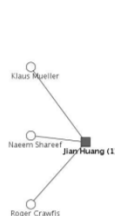
ERI enables low SWaP sensor processing, higher-level SA



Targeting



Entity resolution (ER)

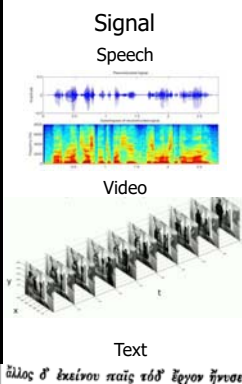


Vertex nomination (VN)

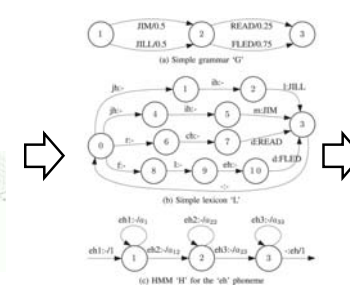


ERI enables targeting updates in real time

Search



Finite state transducer model



Recognition

"I have a cat"

Activity:
Walking through hallway

"Greek is hard to learn"

ERI enables search updates in real time



www.darpa.mil

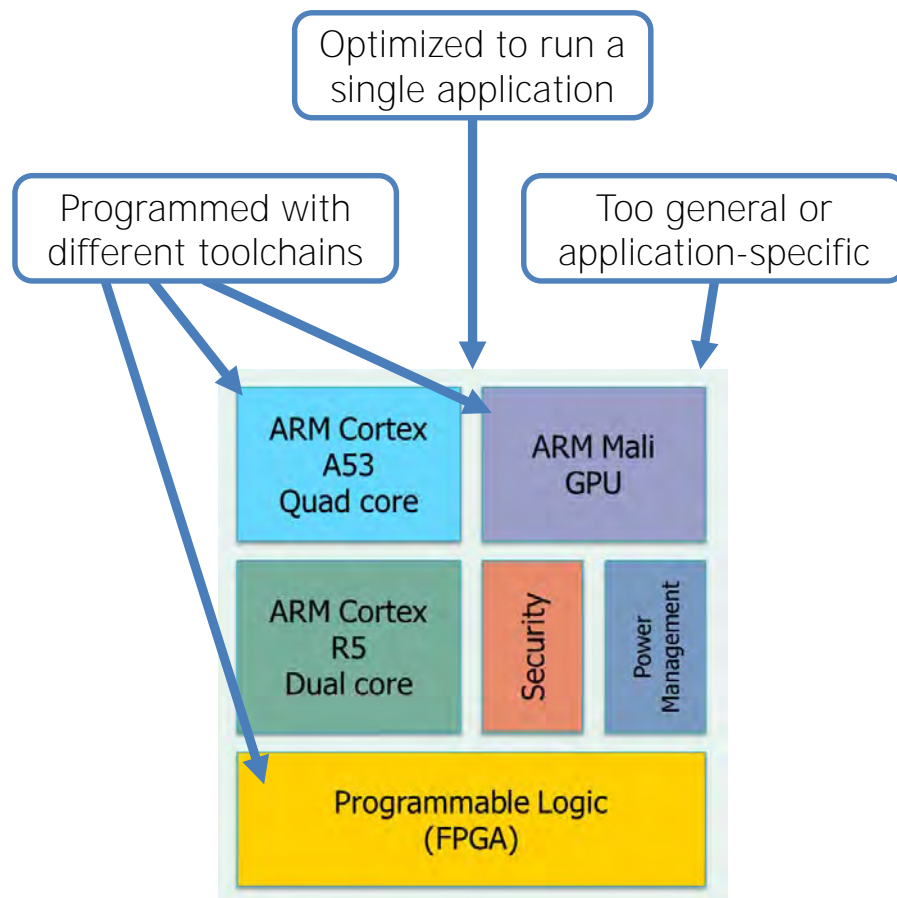
Domain-Specific System on Chip (DSSoC)

Tom Rondeau
DARPA/MTO

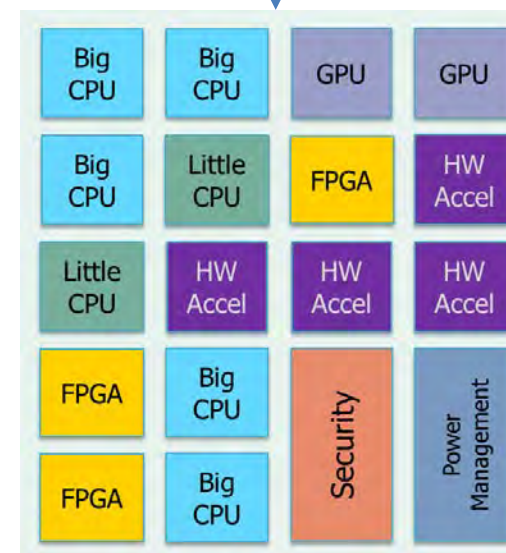
ERI Proposers Day

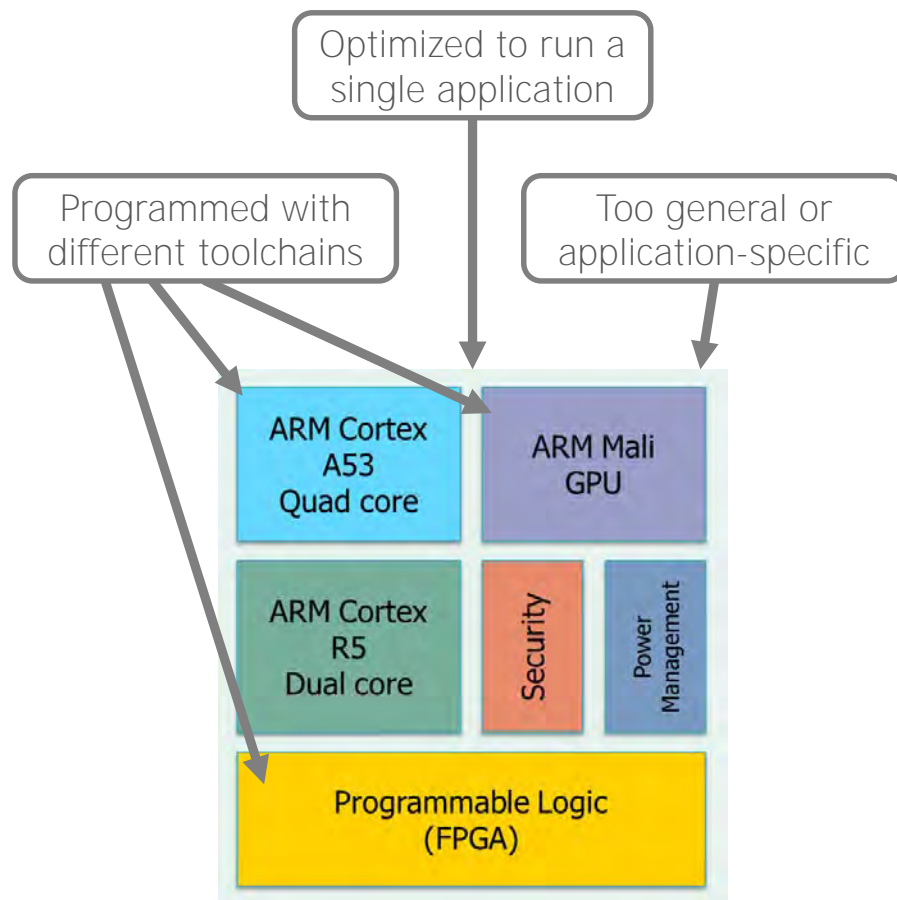
12/19/2018



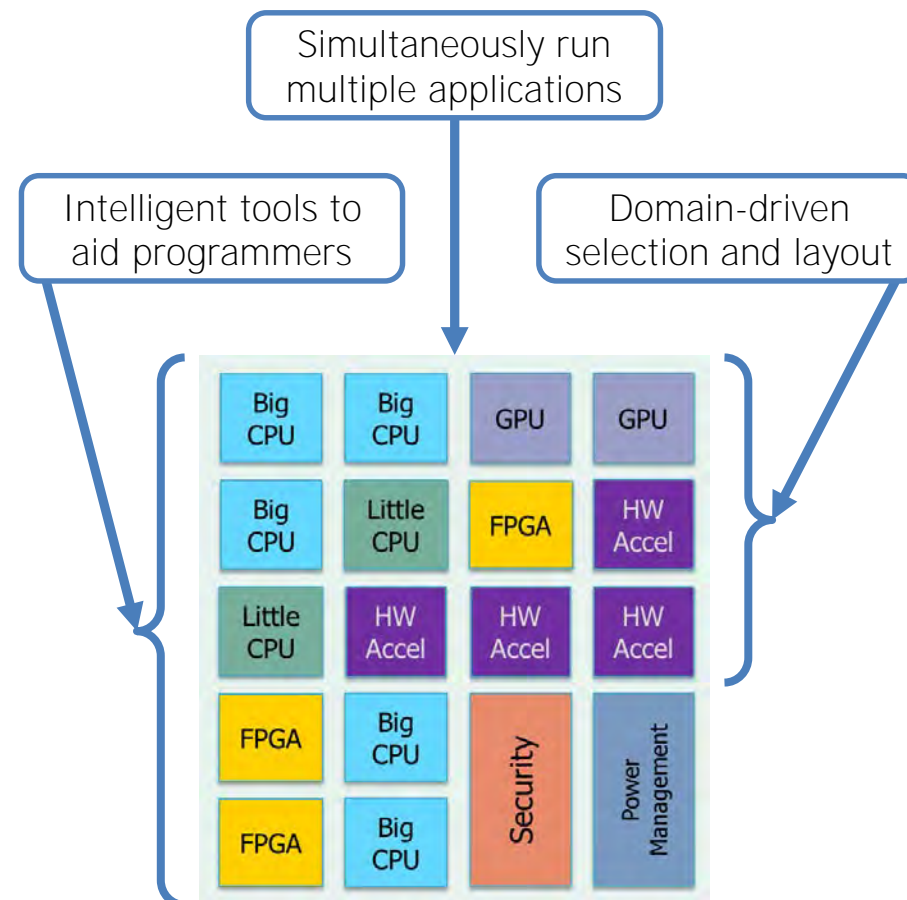


Increasing integration of specialized computing cores
-- And no idea how to use them!





We cannot fully utilize our current multi-processor systems



We want to manage an increase of capabilities with integrated specialized accelerators

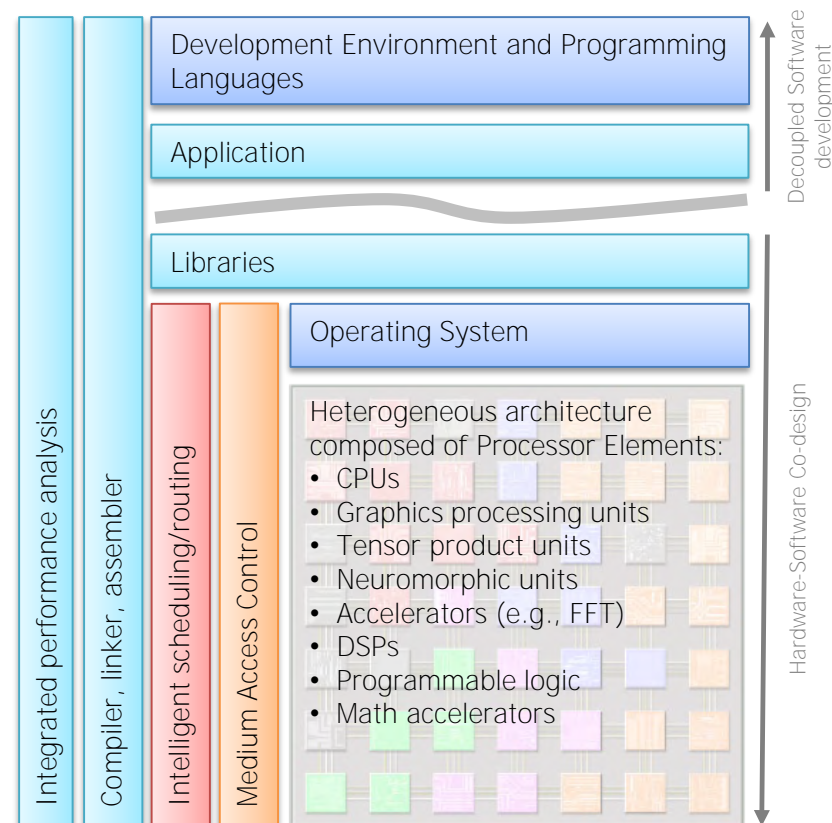
Three Optimization Areas

1. Design time
2. Run time
3. Compile time

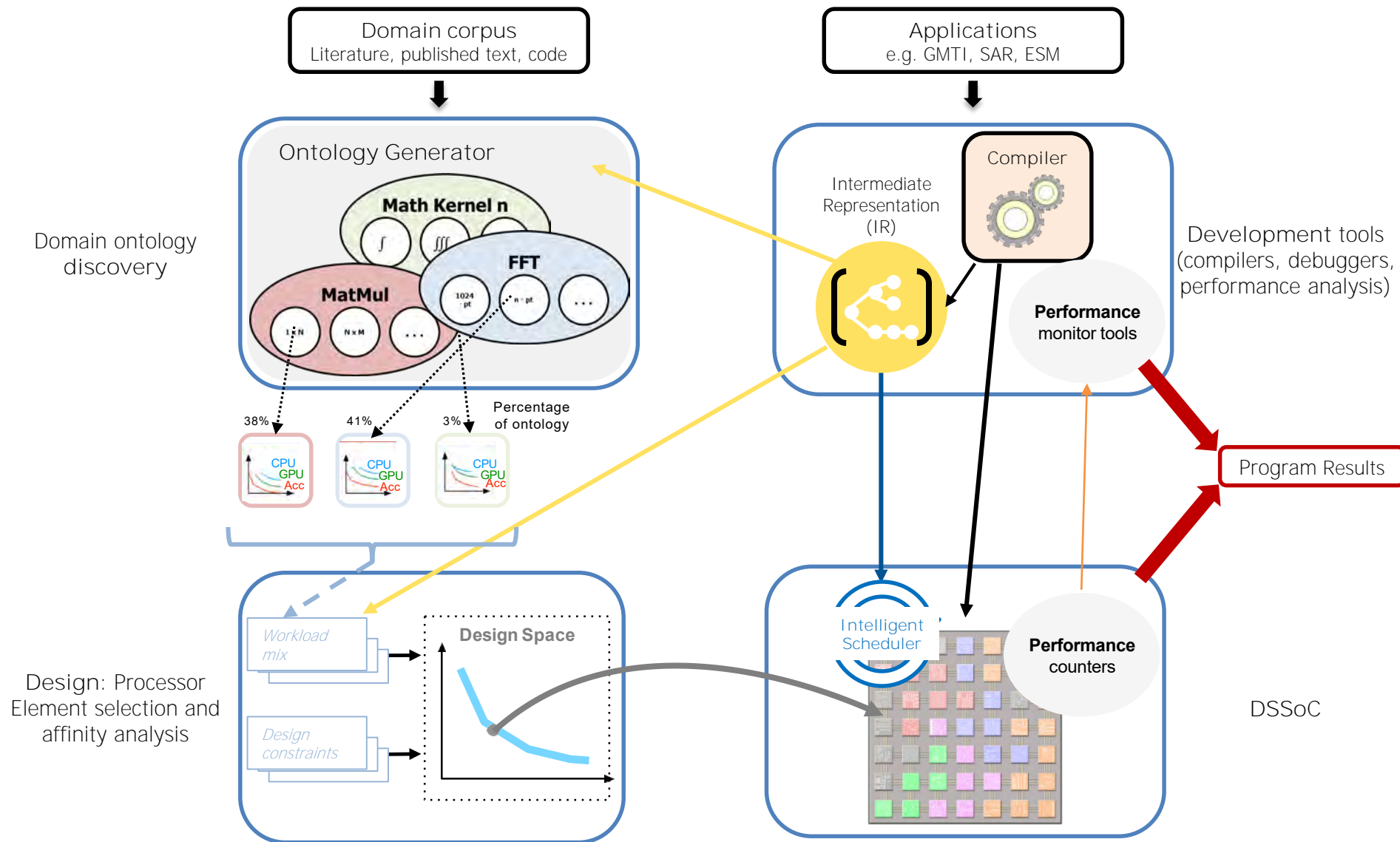
Addressed via five program areas

1. Intelligent scheduling
2. Domain representations
3. Software
4. Medium access control (MAC)
5. Hardware integration

DSSoC's Full-Stack Integration

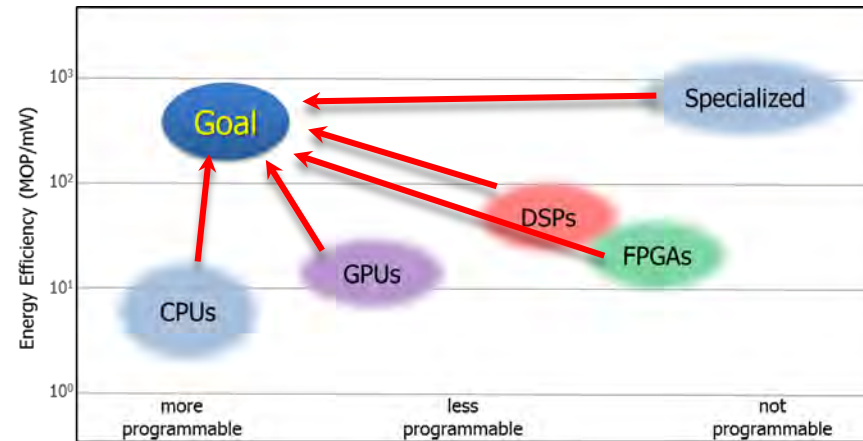
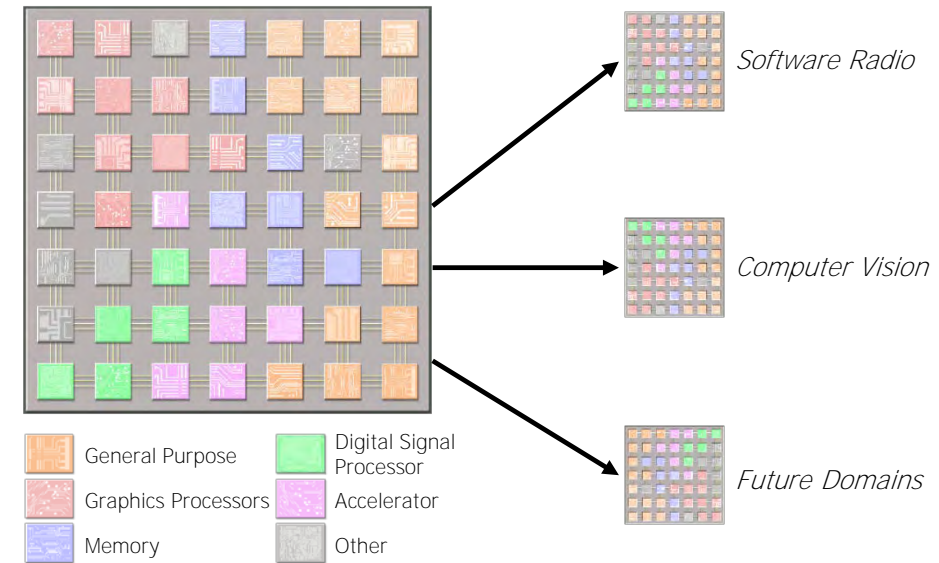


Looking at how Hardware/Software co-design is an enabler for efficient use of processing power

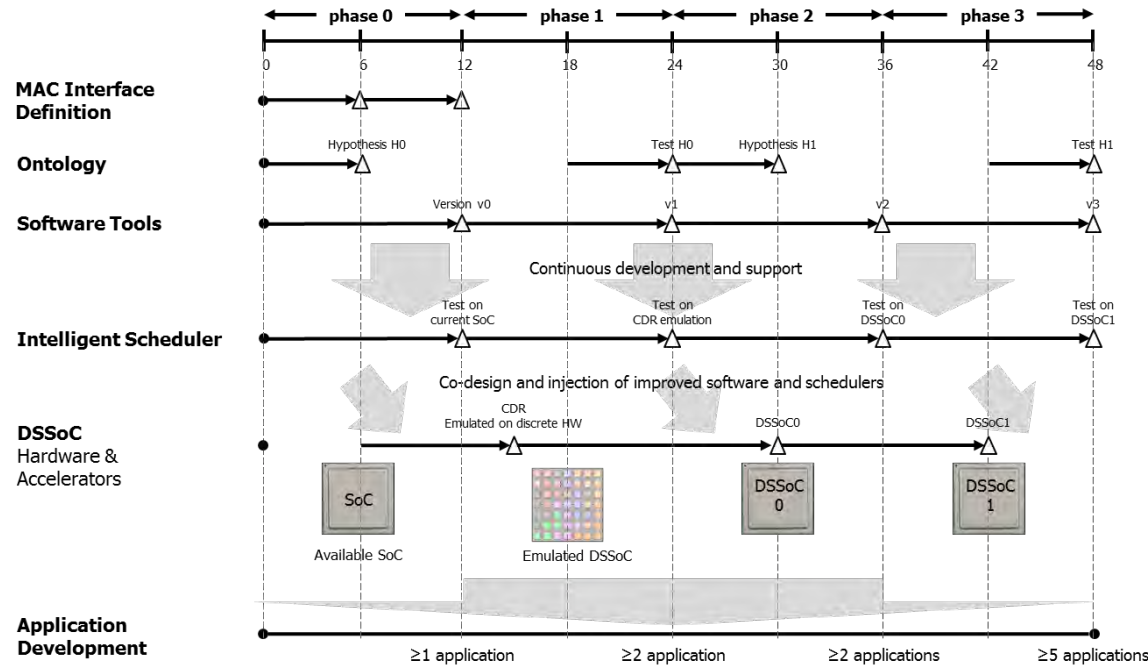




DSSoC: More of a Software Program than a Hardware Program



DSSoC will enable rapid development of multi-application, heterogeneous systems through a single programmable device



DSSoC Performers

Arizona State University (SDR/comms)

IBM (autonomous vehicles)

Oak Ridge National Laboratory (SDR)

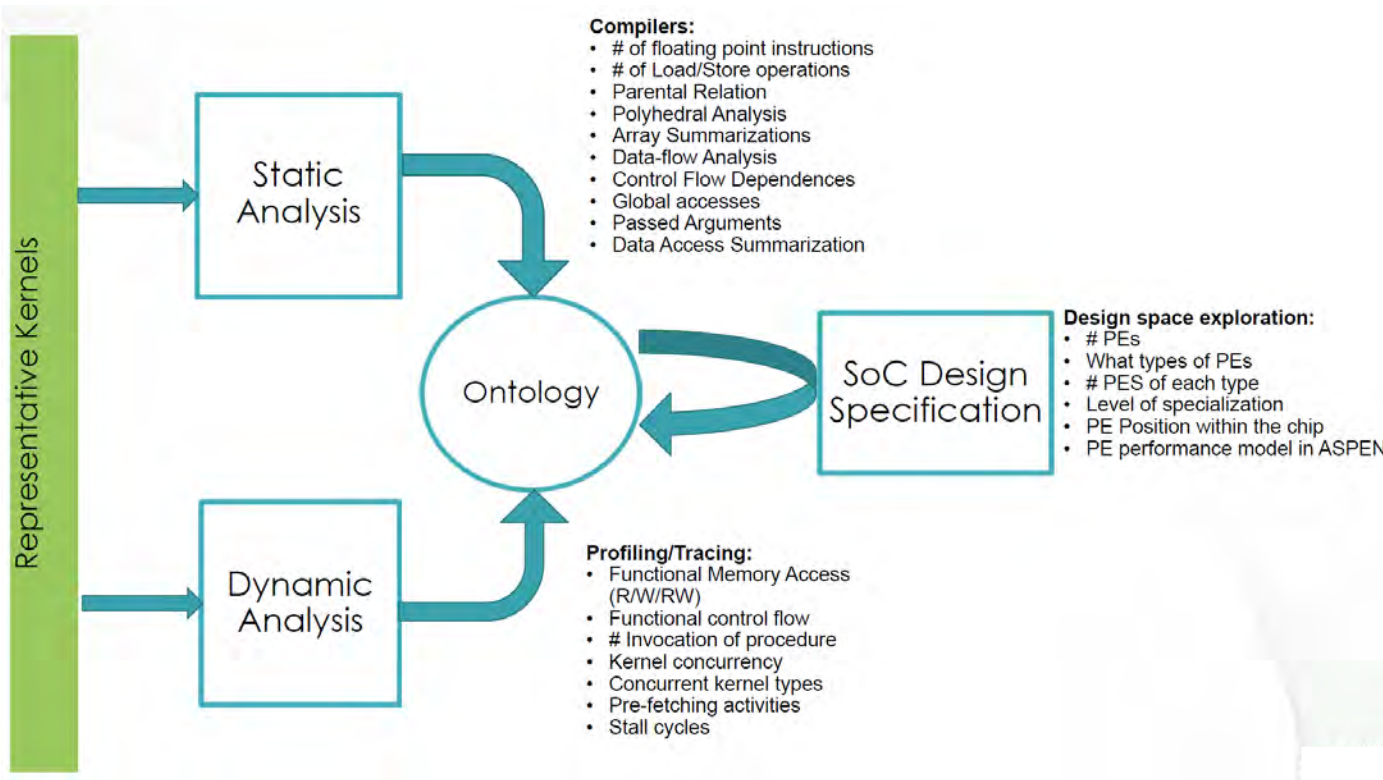
Stanford University (computer vision)

Raytheon (SDR/adaptive beamforming)

	Phase 1	Phase 2	Phase 3
Chip & Scheduler			
Number of simultaneous apps	≥2	≥2	≥5
Integration time for new accelerators ¹		≤3 months	≤3 months
Power savings relative to previous phase		≤80% ²	≤80% ³
Utilization of PEs ⁴	≥80%		≥90%
Max. time per scheduler decision	≤500 ns	≤50 ns	≤5 ns
Medium Access Control (MAC)			
Latency (PE to PE)	≤500 ns	≤50 ns	≤5 ns
Throughput (PE to PE)	≥25 Gbps	≥50 Gbps	≥100 Gbps
Power	≤50% of chip	≤40% of chip	≤20% of chip

Power Constraints	
Embedded System (cell phone)	≤ 5 W
Portable System (laptop)	≤ 25 W

1. Three months to integrate new accelerators into DSSoC; enforced by program timeline
2. Compare the intelligent scheduler on DSSoC0 to the intelligent scheduler controlling the commercial SoC from phase 0.
3. Compare the intelligent scheduler on DSSoC1 to the intelligent scheduler on DSSoC0.
4. Ontology explains the required PEs and utilization; measure average utilization over developed apps.

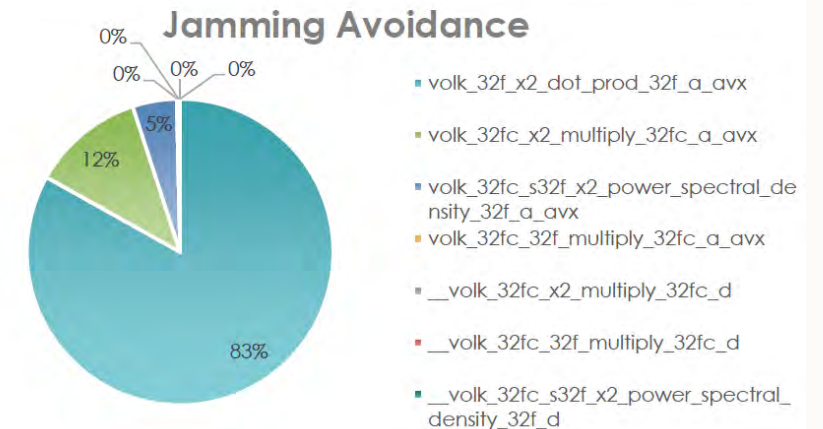
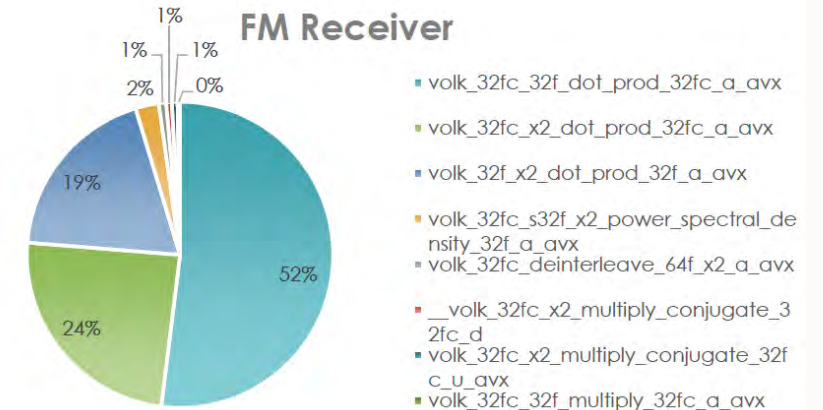


Discover kernels and quantitative characteristics

- **Static Analysis**
 - Architecture-independent features
 - Structural knowledge for the kernels
 - Similarity relationships between kernels
- **Dynamic Analysis**
 - Determine type & quantity of kernels executed
 - Facilitate collecting dynamic system information
 - Power consumption measurement/estimation

Discovering what we need to accelerate in hardware

Example SDR application analysis





Stanford: Tape Out Early and Tape Out Often

Hardware design process should capture more than just RTL

1. Logical Design
2. Physical Design
3. Testing/Verification
4. Software API

Cloud-scale video data mining

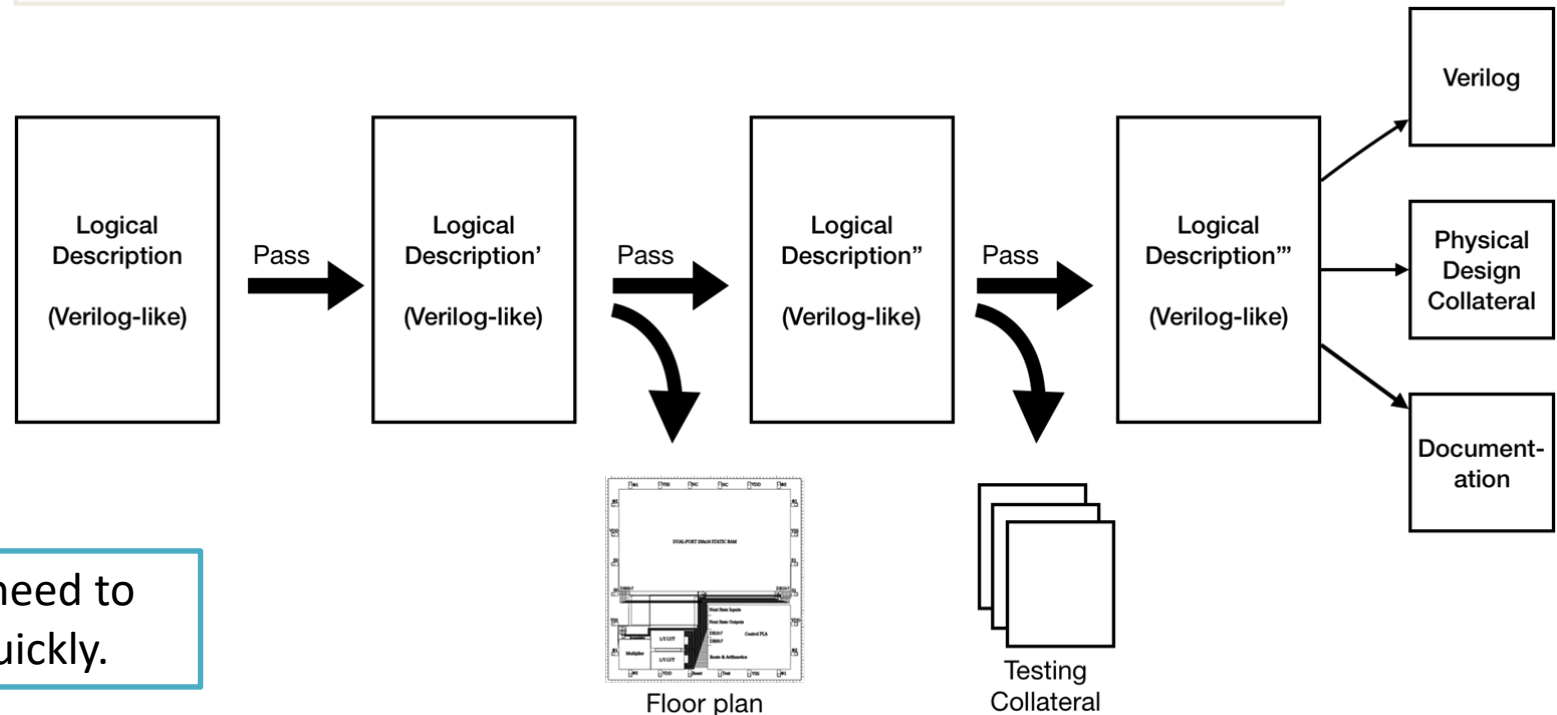
- Image processing and DNN inference on millions of hours of video
- Offline training on video collections

Real-time video stream (& multi-stream) processing

- Computational photography, autonomous vehicles, VR

Low latency, ultra-low power deployments

- Always on sensing/wakeboarding, sense/process/display

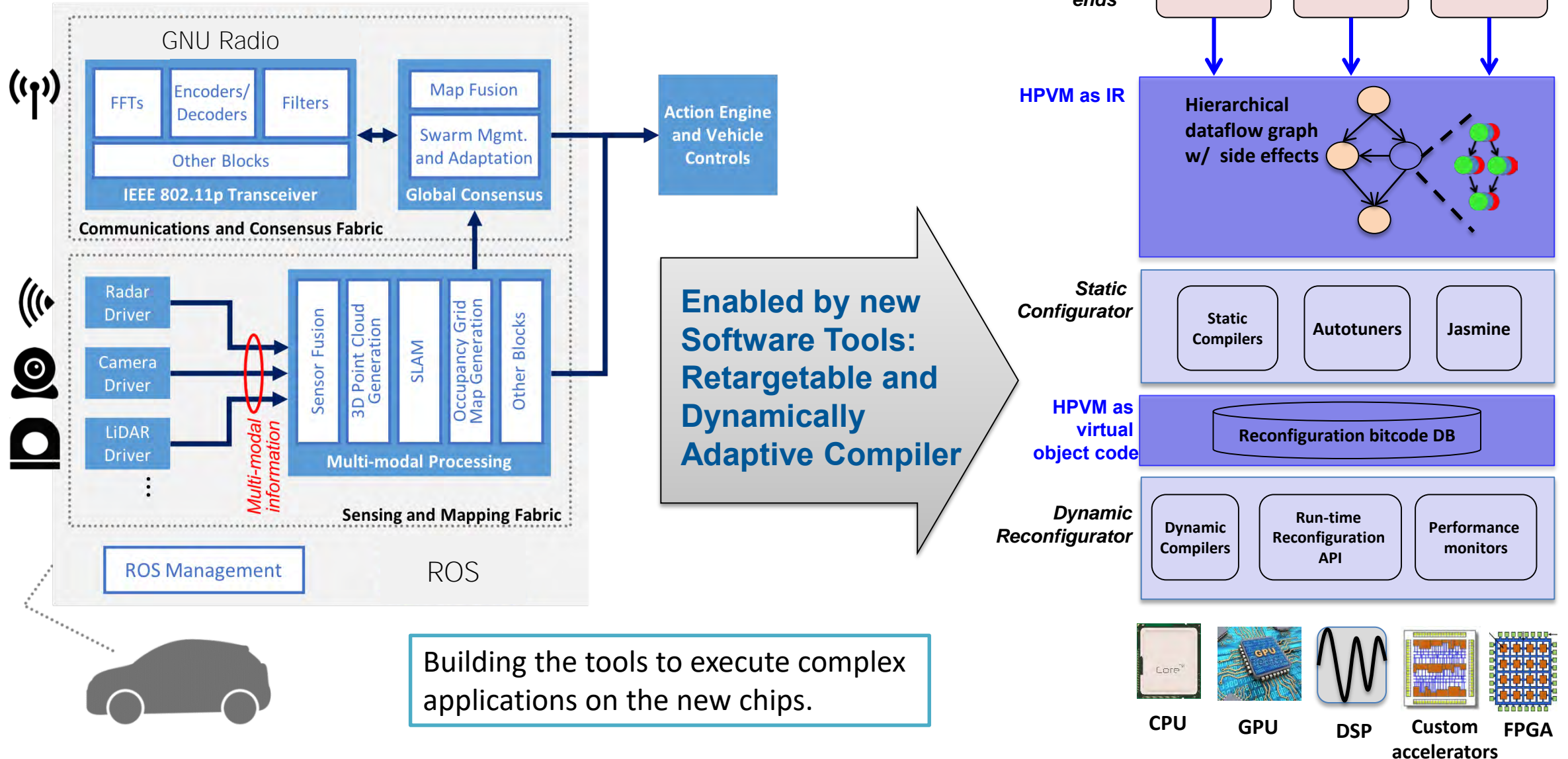


Once we know what to build, we need to construct the chip and lay it out quickly.



IBM: Cooperative Connected Vehicles Application

Open-source: <https://github.com/IBM/era>

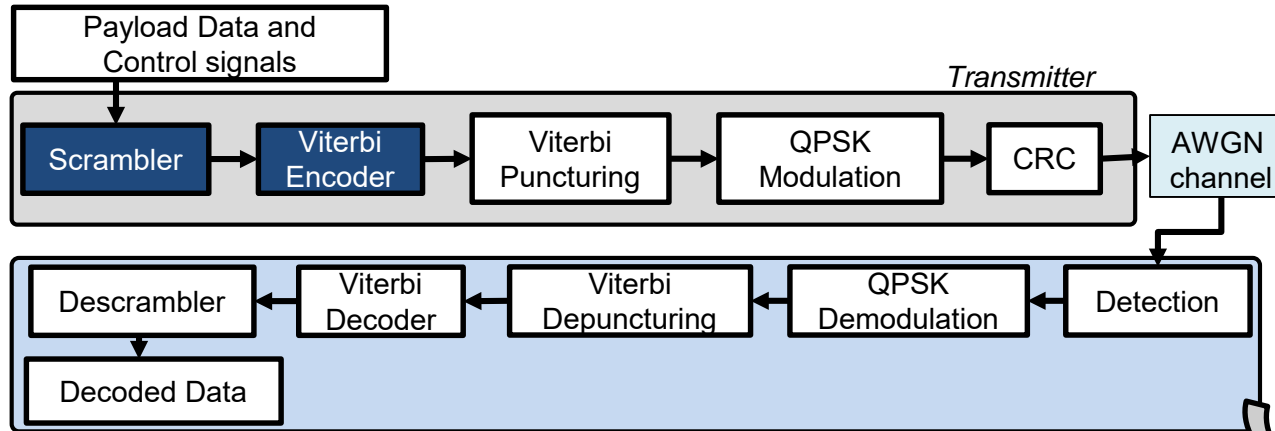




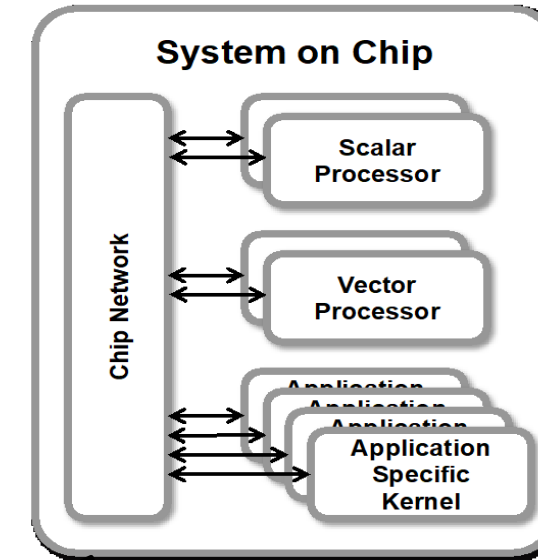
Arizona State: Intelligent Scheduler for Runtime Optimization

Software Radio applications

- Remove need for expert hand-tuning of function blocks
- Learn to schedule multiple applications across all processor elements



Runtime optimization and mapping of applications onto the underlying processor elements.



Adaptive Interference-Mitigation Communications

- Bandwidth = 10 MHz $\rightarrow t_{\text{chip}} = 0.1 \mu\text{s}$
- Coherence Interval \sim few ms $\rightarrow t_{\text{waveform}} = 1 \text{ ms}$
- Delay Spread $\sim 500 \text{ ns} \rightarrow n_{\text{taps}} \sim 20$
- INR = 30 dB, SNR = 10 dB $\rightarrow n_{\text{pilot}} > 1000 \text{ chips}$
- Oversampling = 2

Advanced radio algorithms have huge computational burdens

$$\mathcal{O} \sim k (n_{\text{ant}} \cdot n_{\text{tap}} \cdot n_{\text{oversamp}})^2 n_{\text{pilot}} n_{\text{oversamp}} / t_{\text{waveform}} \\ \approx 1 \text{ TOP/s}$$

Space-time beamformer:

$$\mathbf{w} = \hat{\mathbf{R}}^{-1} \mathbf{v} = (\tilde{\mathbf{Z}} \tilde{\mathbf{Z}}^{\dagger})^{-1} \tilde{\mathbf{Z}} \mathbf{u}_{\text{pilot}}^{\dagger} \in \mathbb{C}^{(n_{\text{ant}} \cdot n_{\text{tap}} \cdot n_{\text{oversamp}}) \times 1}$$



Phased Array ADBF	SAR Radar	Range Doppler Radar Search	Sat Comms	EW Transceiver
Correlation (Matrix Inverse)	Clutter Suppression (Matrix Inverse)	Auto-Correlation (FIR Filter)	Encoder (FIR Filter)	Pulse Intercept
Adaptive Beamform (Matrix Multiply)	Conventional Beamform (Matrix Multiply)	Target Detection	High Speed Modem	Low Latency DRFM
Null Creation	1M Pt FFT	1000 Pt FFT		64 Point FFT

Applying these ideas to hard problems in hostile environments

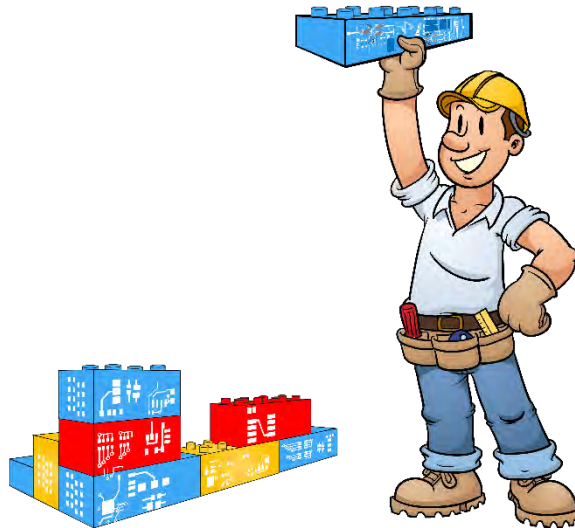


www.darpa.mil

Circuit Realization at Faster Timescales (CRAFT)

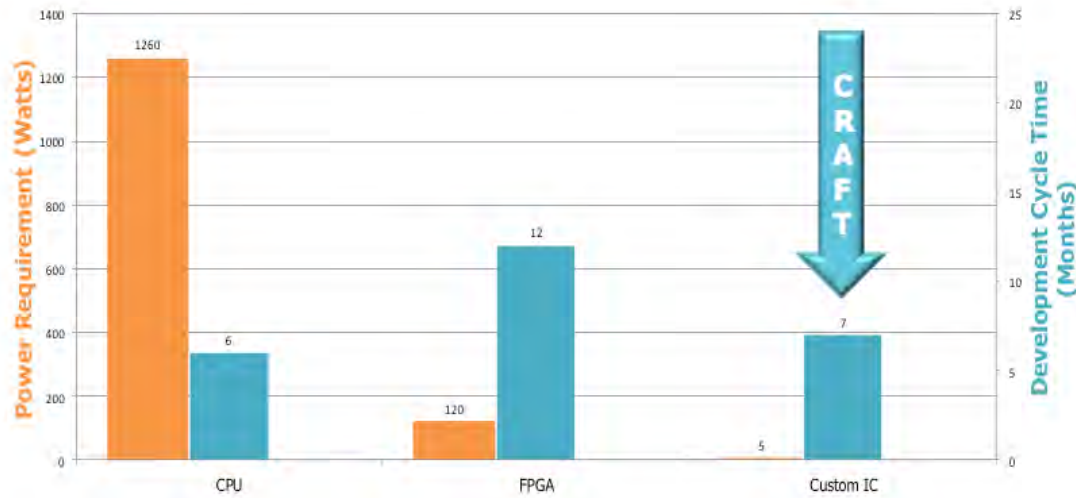
PM: Dr. Linton Salmon

Presenter: Mr. Andreas Olofsson

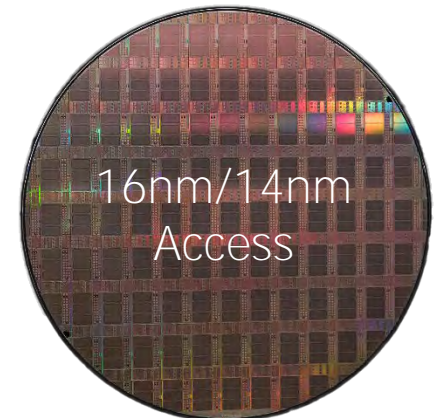


December 19, 2018

To sharply reduce the barriers to DoD use of custom integrated circuits built using leading-edge CMOS technology while maintaining the high level of performance at power promised by this technology.



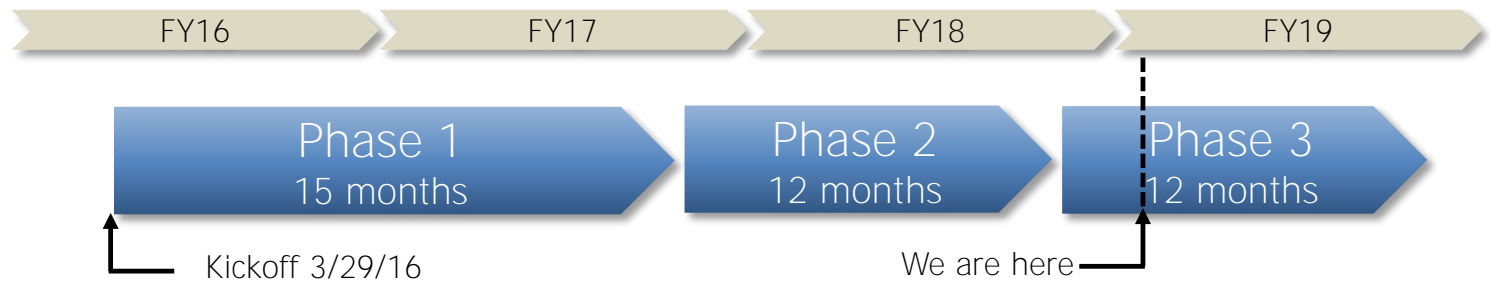
Efficiency of Design (BAA 15-55)



Efficiency of Access (MPW Runs)

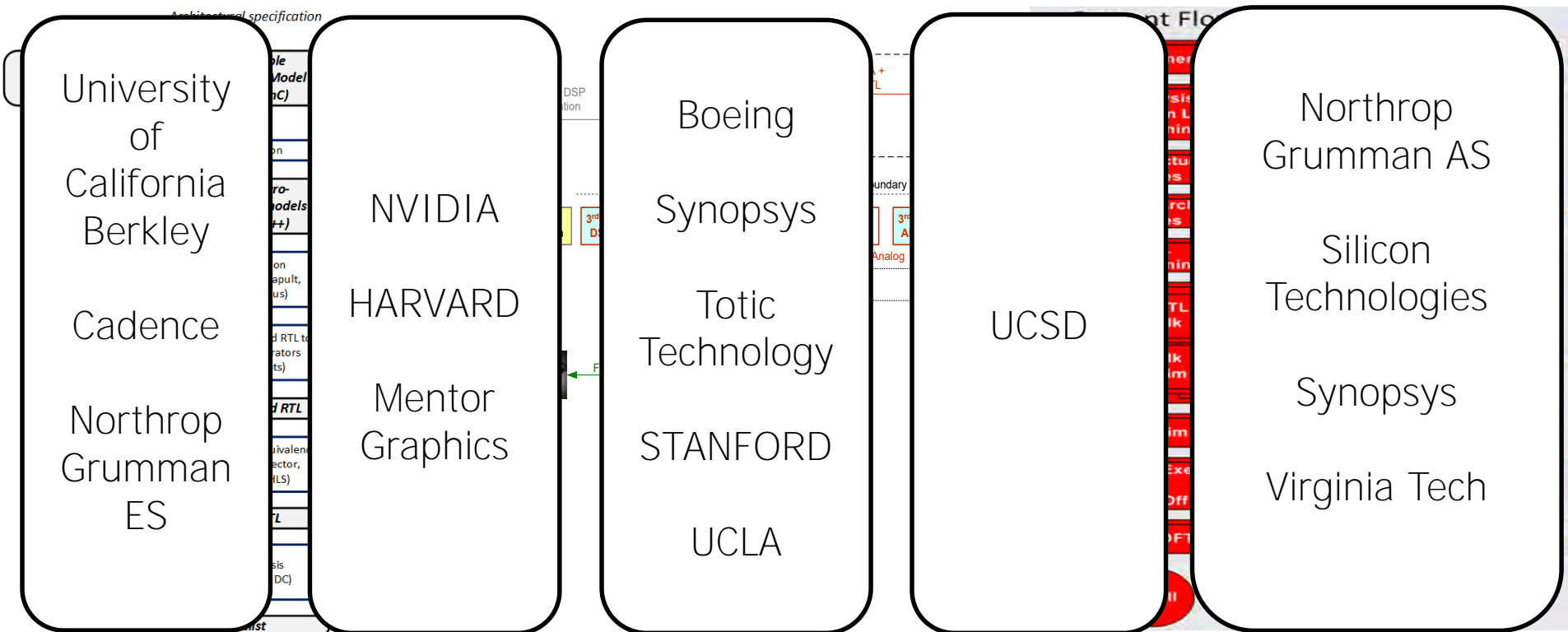


CRAFT will enable more efficient custom IC design/fabrication that will enable HIGH performance electronic solutions FASTER and with more FLEXIBILITY



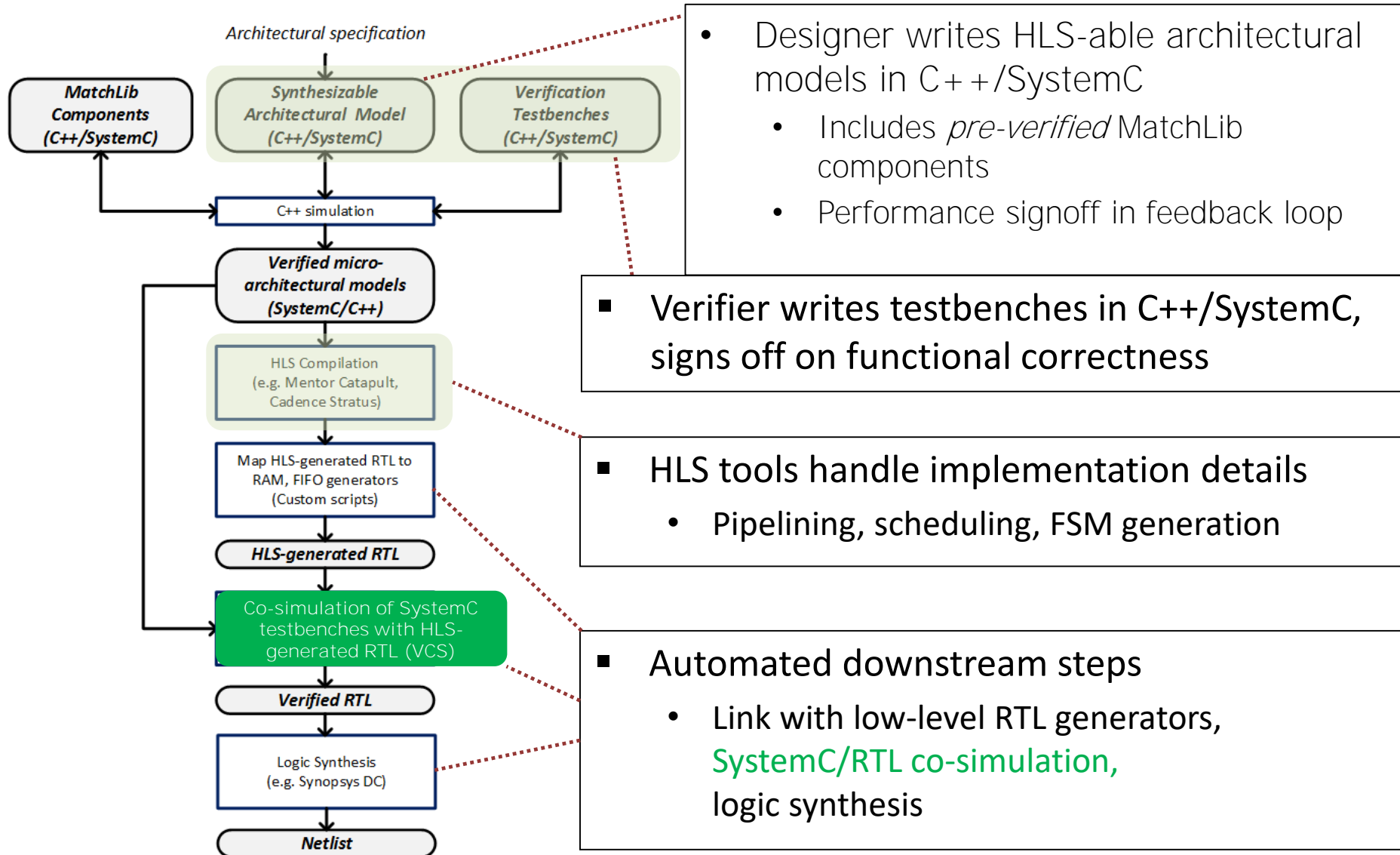
-
- TA – 1: Design CRAFT has developed new design flows that reduce custom IC design cycle time by **10X**
-
- TA – 2: Port CRAFT uses the new CRAFT design flows to reduce the effort to re-design an SoC to a new process by **5X**
-
- TA – 3: Vault CRAFT establishes a unified design environment and repository to enable IP reuse by DoD performers
-
- Access CRAFT provides access to 16nm FinFET technology through 4 dedicated, DARPA rapid MPW runs
-

- Design, integration and verification of blocks at architecture level
 - Applied to both Analog and Digital designs
- Faster integration of third party IP
- Facilitate small design teams for IC design projects

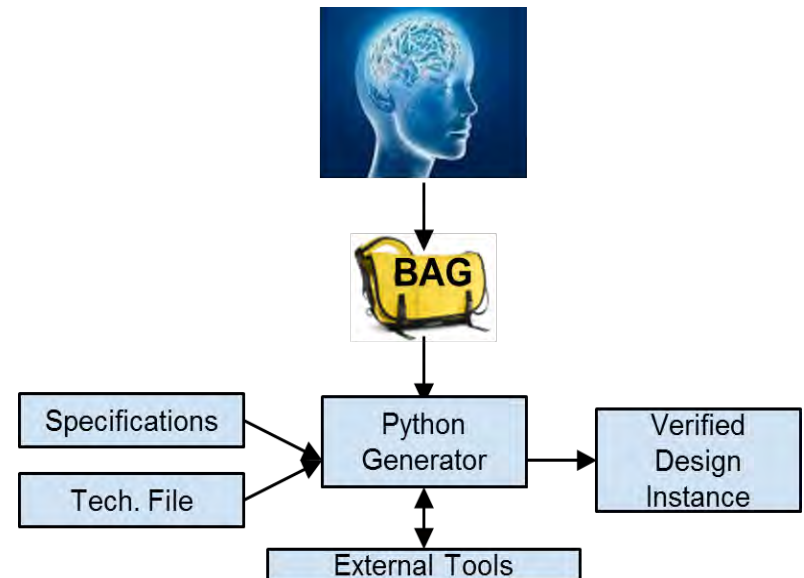
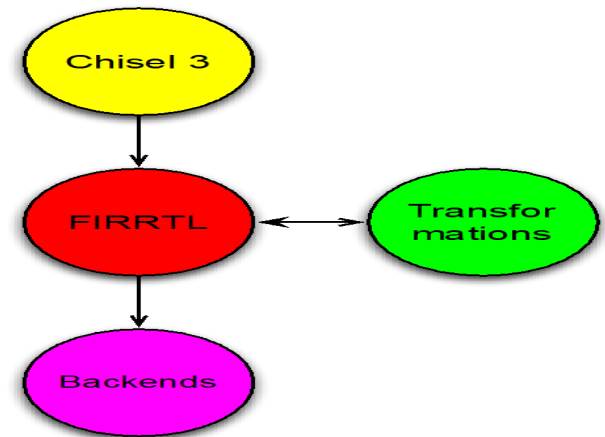




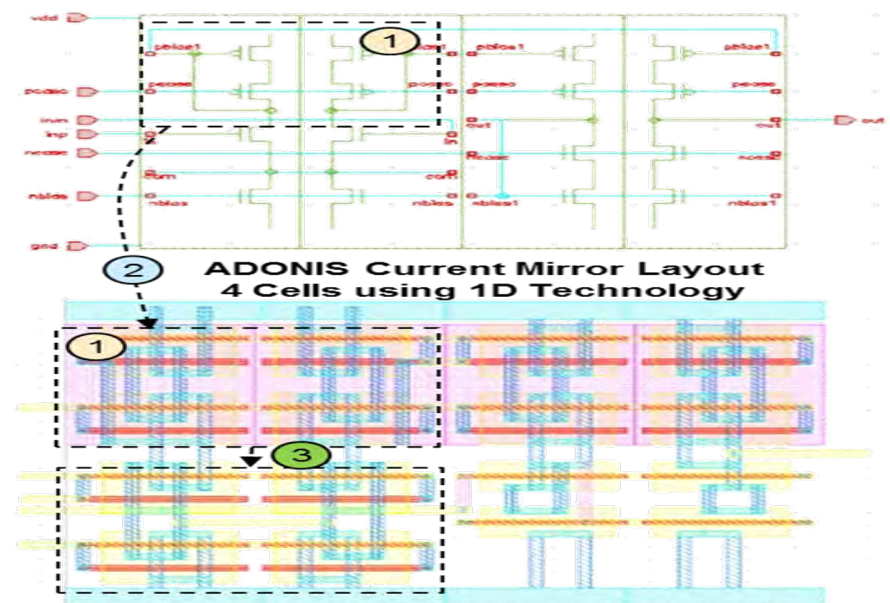
Advances in Digital Front-End and Verification (NVIDIA)



- Hardware construction language
 - Writes Generators that construct hardware
 - Generators create RTL and associated tests to verify the design
 - Allows greater reuse of IP
- Codifies analog designer's methodology
- Python based framework for capturing design specification and layout procedure
- Produces HDL compatible functional model for full SOC verification
- Allows subsequent designs to be done faster



- Automatic Placement for 100% accurate schematic vs. layout
 - Primitive cells are developed from technology file.
 - Schematic is drawn based on primitive cells.
 - Optimized interconnect is routed based on correct loading between cells
- Libraries can be migrated to target technology fast.





Results to Date - Design

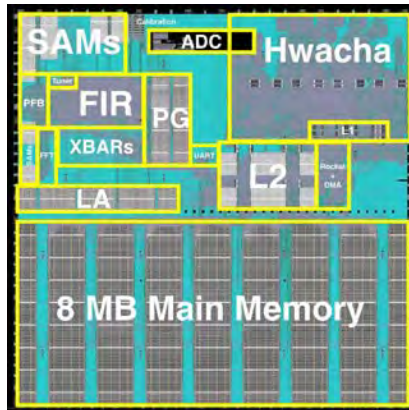
Metric	Current Best in Class	End of Program Goal	Phase 2 Results		
			Nvidia	UCB	STI
Average Digital	3.5 kgates/eng-day	100 kgates/eng-day	41 kgates/eng-day	53 kgates/eng-day	NA
Peak Digital			120 kgates/eng-day	180 kgates/eng-day	NA
Average Analog	.2 blocks/eng-week	1.5 blocks/eng-week	NA	2.6 blocks/eng-week	6.3 blocks/eng-week
ADC Design	40 eng-weeks	4 eng-weeks	NA	4 eng-weeks	5.8 eng-weeks
Overall Design	100%	10%	9%	12%	11%

Representative SoC Complexity

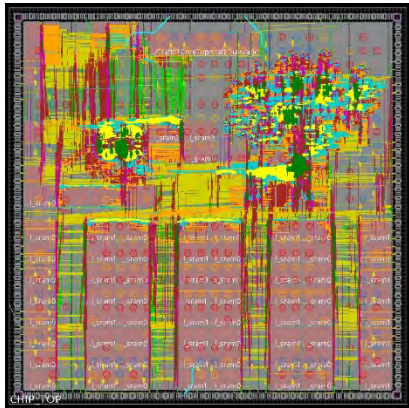
SoC Design	Digital Modules	Logic Size	Memory Size	Mixed-Signal	Die Size
UCB	8- Rocket Cores 8-Vector Processor	24.6M gates	29Mb	26Gb/s SERDES 8-bit ADC	25mm ²
Nvidia	DNN PE NoP Router	7.6M gates	6.4Mb	NA	6mm ²

Large increases in digital and analog design efficiency demonstrated on moderately complex SoCs

TSMC 16FF

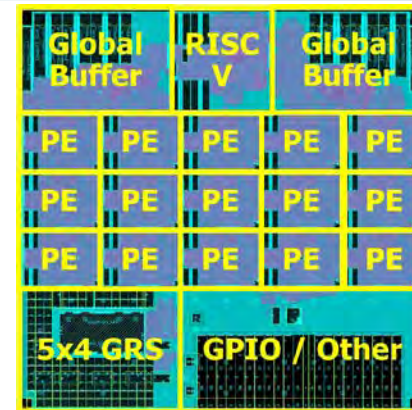


GF 14LPPXL

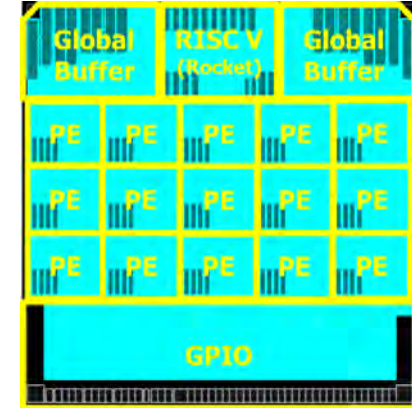


UC-Berkeley DSP SoC

- Original approach estimate:
- Phase 1 hours: 14,000
- Phase 2 achievement: 2,663 hours
- Percentage: 19%



TSMC 16FF

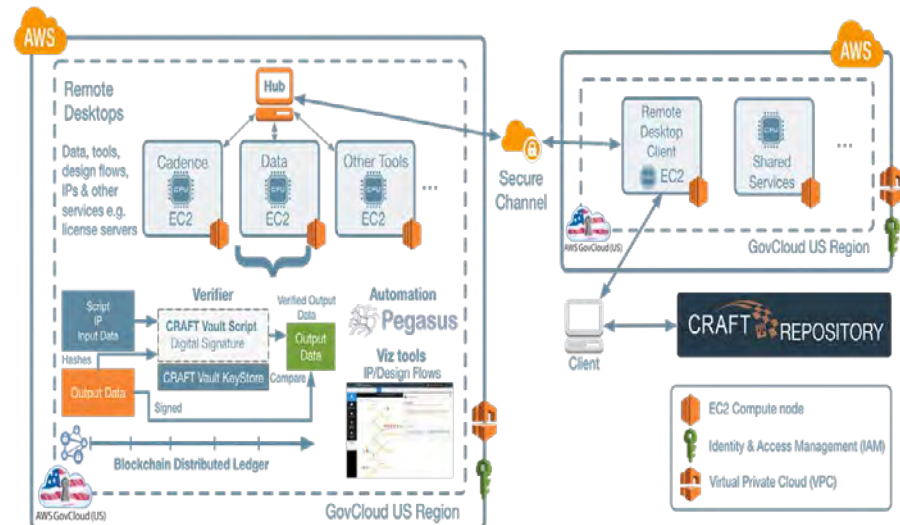


GF 14LPPXL

Nvidia RC17 ML SoC

- Original approach estimate:
- Phase 1 hours: 3216 hours
- Phase 2 achievement: 754 hours
- Percentage: 23%

- Developing Space deployable processor using CRAFT digital design flow
- Developing variations of the Space deployable processor architecture in 16nm and 45nm PD-SOI process technology utilizing RISC-V processor
- Secure collaborative environment for CRAFT Design Teams for ASIC design
- Management of authentication and authorization privileges
 - Tools, IP, and user designs and data
- Automated access to virtualized secure environments via GovCloud AWS
- Ready-to-use pre-tested deployment of entire design flow, i.e., all tools and their dependencies





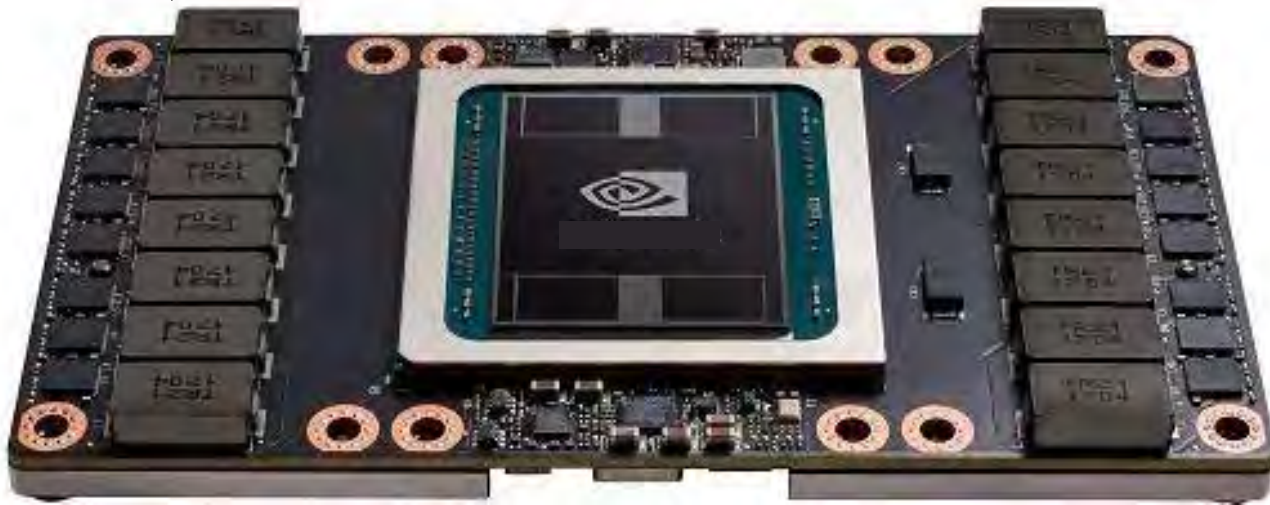
www.darpa.mil

An Introduction to DARPA's Silicon Compiler Effort

Andreas Olofsson
Program Manager
DARPA/MTO

ERI Defense Applications Proposers Day
Arlington, VA
12/19/2018





- NVIDIA V100 (2017-)
- 0.012um feature size
- 21,000,000,000 transistors

Death by a million papercuts...functional correctness, security, safety, reliability, system performance, IP integration, power management, firmware, system integration, wire delays, placement, routing, clocking, signal integrity, process specific design rules, antenna effects, ESD, multi voltage domains, power gating, multi threshold, floor-planning, I/O structures, flip-chip, wirebonding, 2.5D packaging, TSVs, RDL routing, area minimization, routing congestion, on-chip variability, self heating, stress and proximity effects, electro migration, SEUs, signal integrity, power delivery networks, decoupling, modeling, low voltage operations, cooling, scan insertion, BIST, ATPG, STA, yield optimization, static and dynamic power minimization, and all the EDA tools needed to make this work..

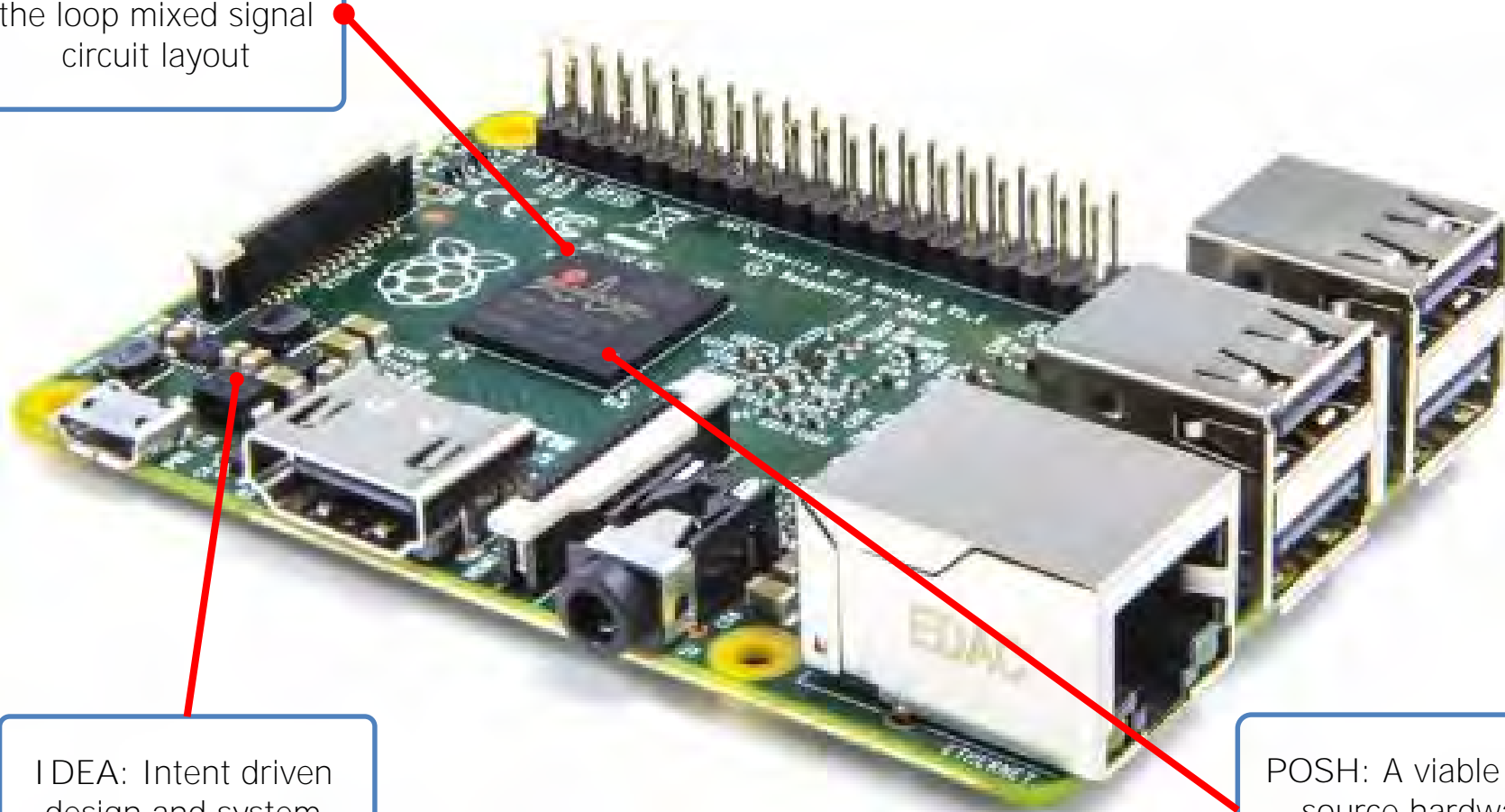


DARPA's \$100M Silicon Compiler Investment

IDEA: No human in the loop mixed signal circuit layout

IDEA: Intent driven design and system synthesis

POSH: A viable open source hardware design ecosystem





```
$ git clone https://github.com/darpa/idea
$ git clone https://github.com/darpa/posh
$ cd posh
$ make soc42
```



2018

- Program Kickoff (Jun)

2019

- First Integration Exercise (Jan)

2019

- Alpha code drop (Jun)

2020

- A usable Silicon Compiler
- 50% PPA

2022

- A great Silicon Compiler
- 100% PPA



What it takes to build a silicon compiler

IDEA Program

11 Teams

16 Subcontractors



26 Professors

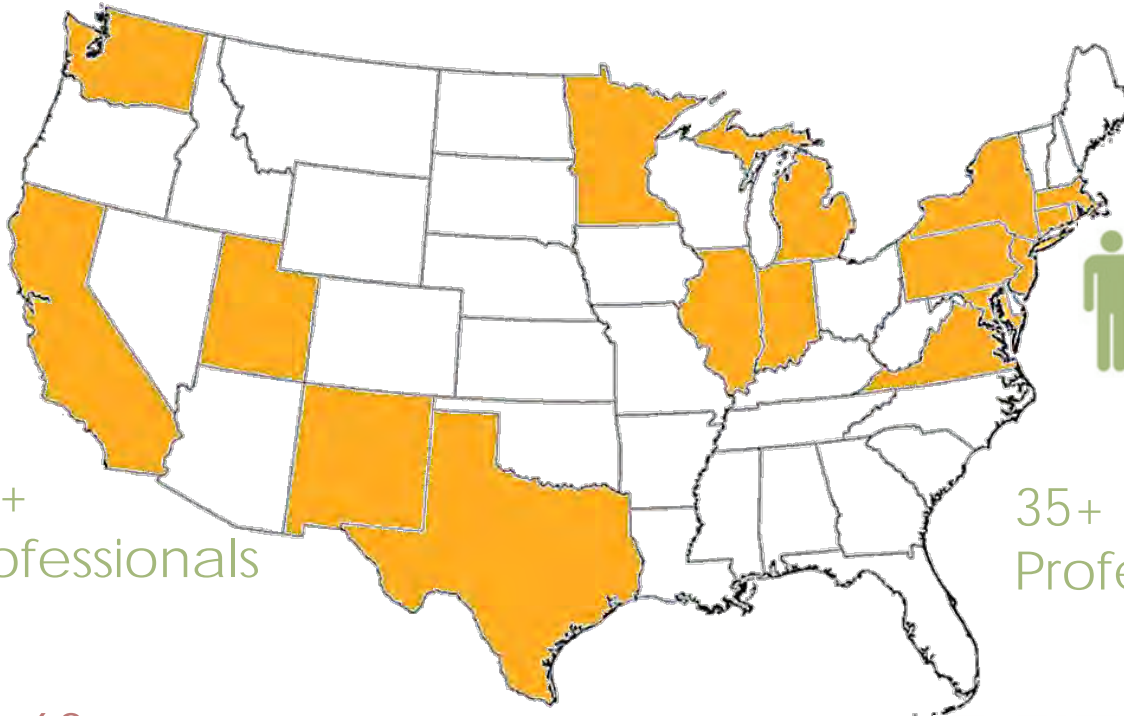


35+ Professionals



60 Graduate Students

22 Teams
Across 17 States



57% Academic Institutions
43% Industry

POSH Program

11 Teams

10 Subcontractors



18 Professors



35+ Professionals



39 Graduate Students



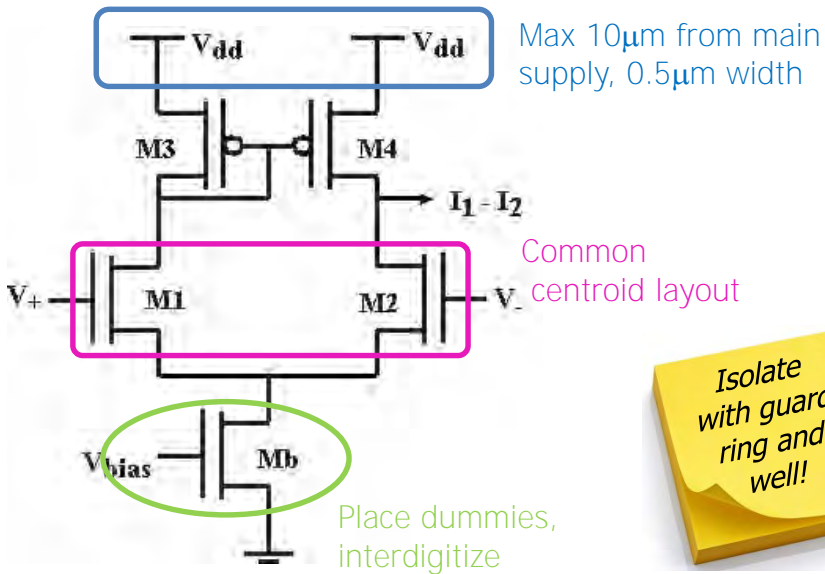
I DEA: Intelligent Design of Electronic Assets



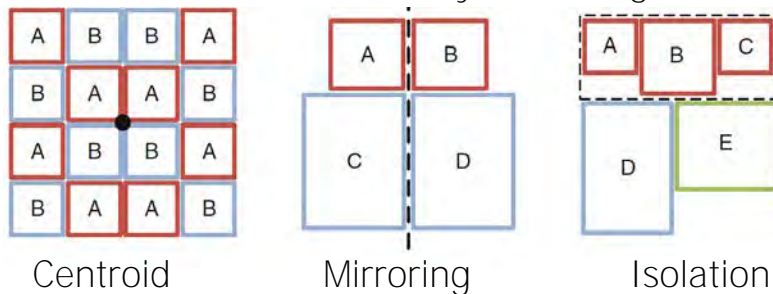
IDEA: No human in the loop layout

Today

Designer provides manual constraints to layout person (or tool)



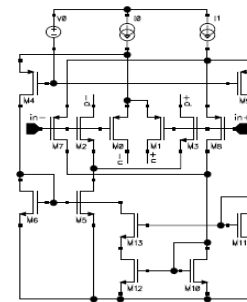
Common Vocabulary of Strategies



Future

Automatically assign constraints based on trained circuit & layout models

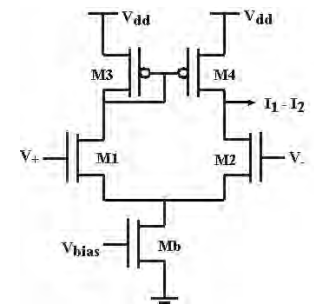
Millions of circuits



Training

Circuit & Layout Models

New circuits



Circuit Classifier

Assign Strategies & Constraints

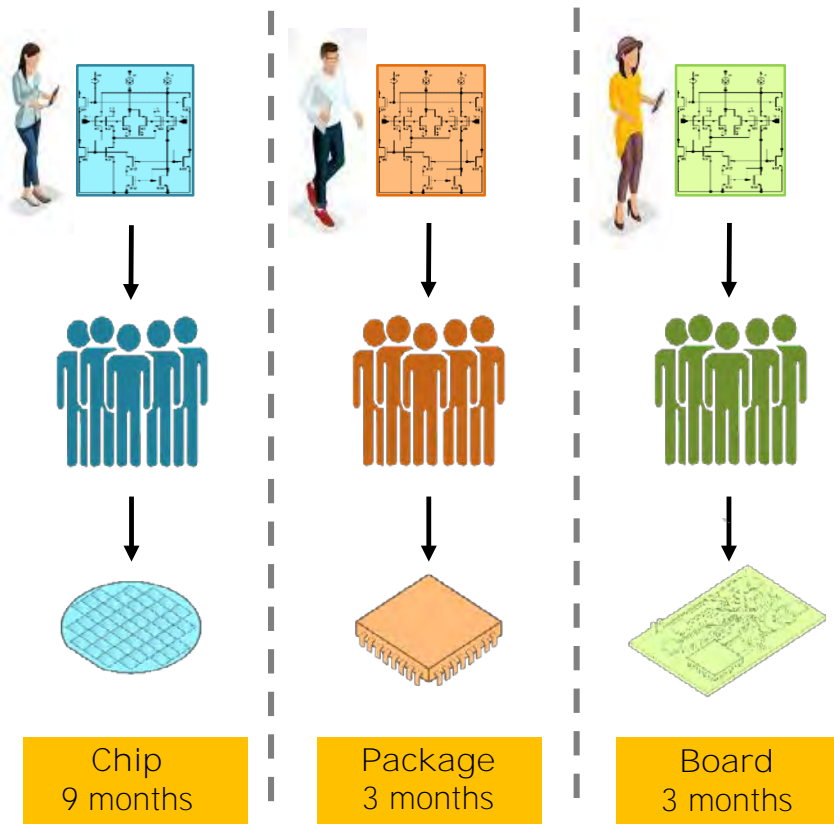
Auto-placement

Auto-routing



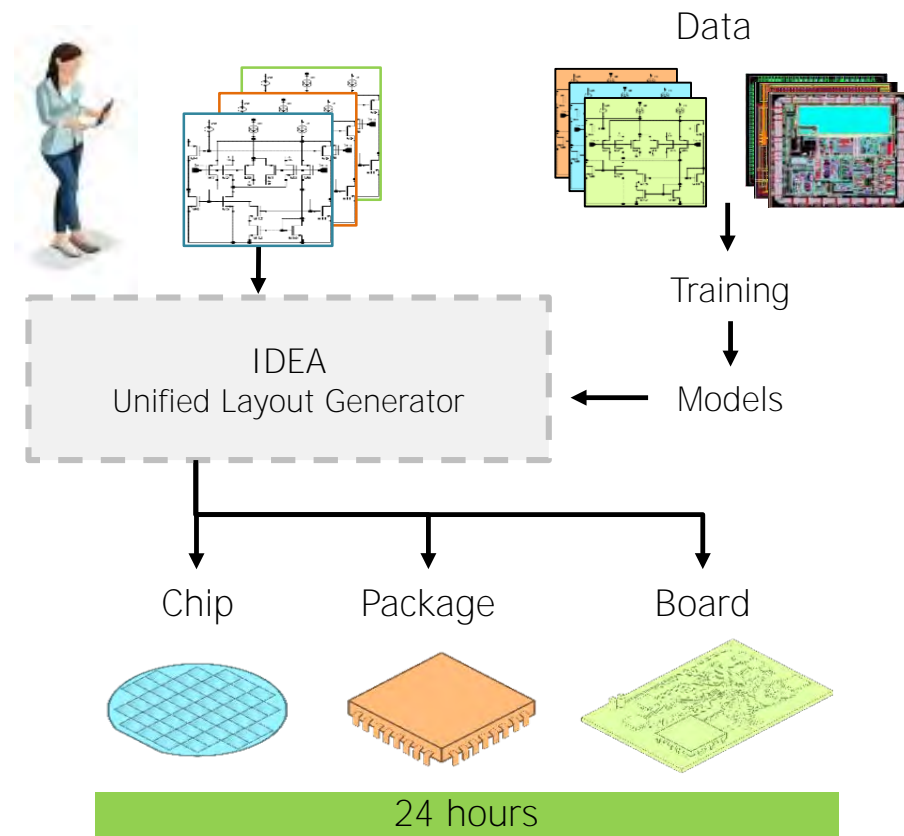
IDEA TA1: A unified electrical circuit layout generator

Today



- Knowledge embedded in humans
- Limited knowledge reuse
- Reliance on scarce resources

Future



- Knowledge embedded in software
- 100% automated hardware compilation
- 24 hour turnaround



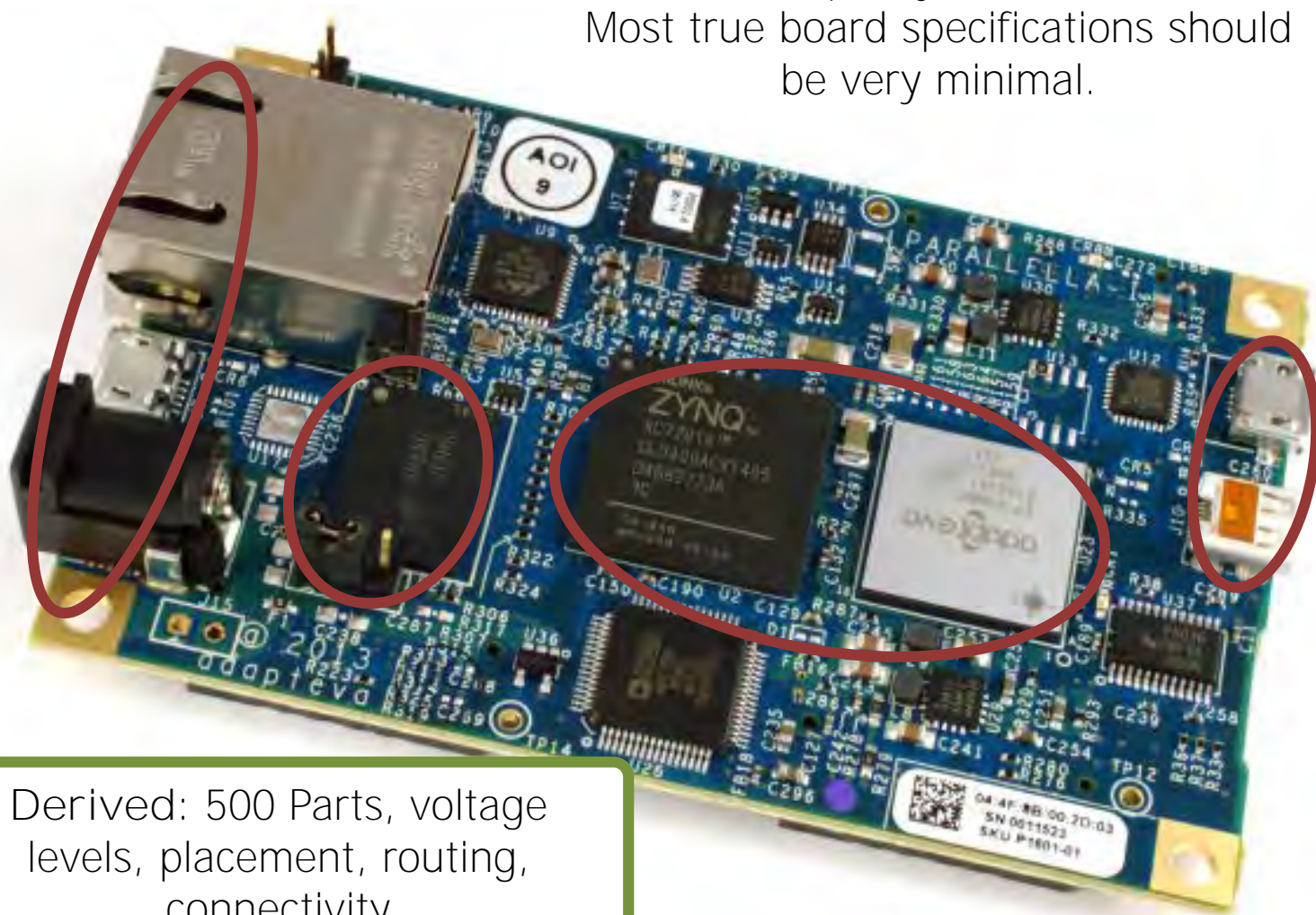
Technical Area	Metrics	Phase 1	Phase 2
IDEA TA-1: Machine Generated Physical Layout	SoC Benchmarks	Government furnished benchmarks 14nm CMOS PDK	Government furnished benchmarks 7nm & 14nm CMOS PDK
	Board Benchmarks	BeagleBone Black ¹	Open Compute Server ²
	SiP Benchmarks	Government furnished benchmarks	Government furnished benchmarks
	Benchmark $PPA_{IDEA}/PPA_{Traditional}^{(3)}$	0.5	1
	Package Complexity	Up to 2 die, 2.5D	Up to 1024 die, 2.5D
	Automation	100%	
	Turnaround time	24 hours	
	Deliverable	Software, license ⁴ , software documentation	



IDEA TA2: Intent Driven Synthesis

Intent: Specify what, not how!
Most true board specifications should be very minimal.

True Specs:
5V
Ethernet
USB
HDMI
1GB RAM
128MB Flash
FPGA
20 GFLOPS
ARM A9

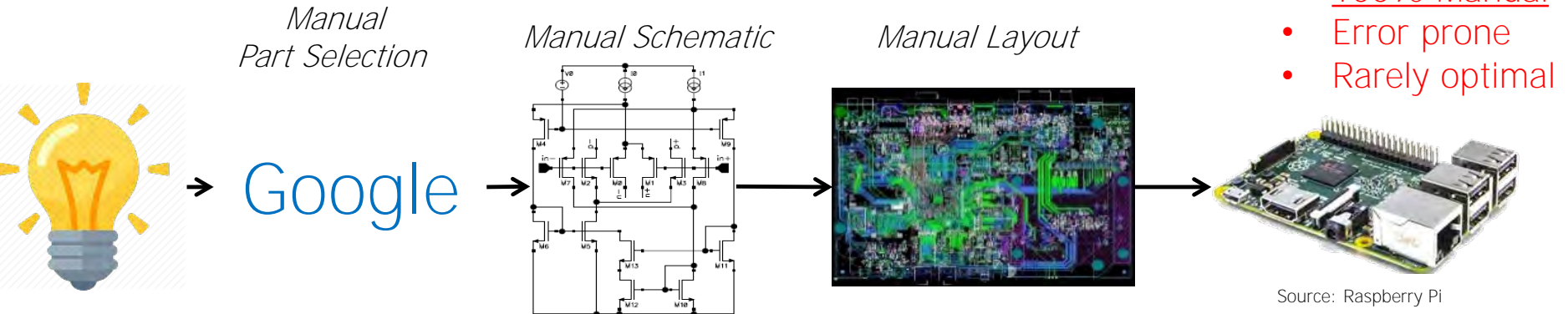


Derived: 500 Parts, voltage levels, placement, routing, connectivity

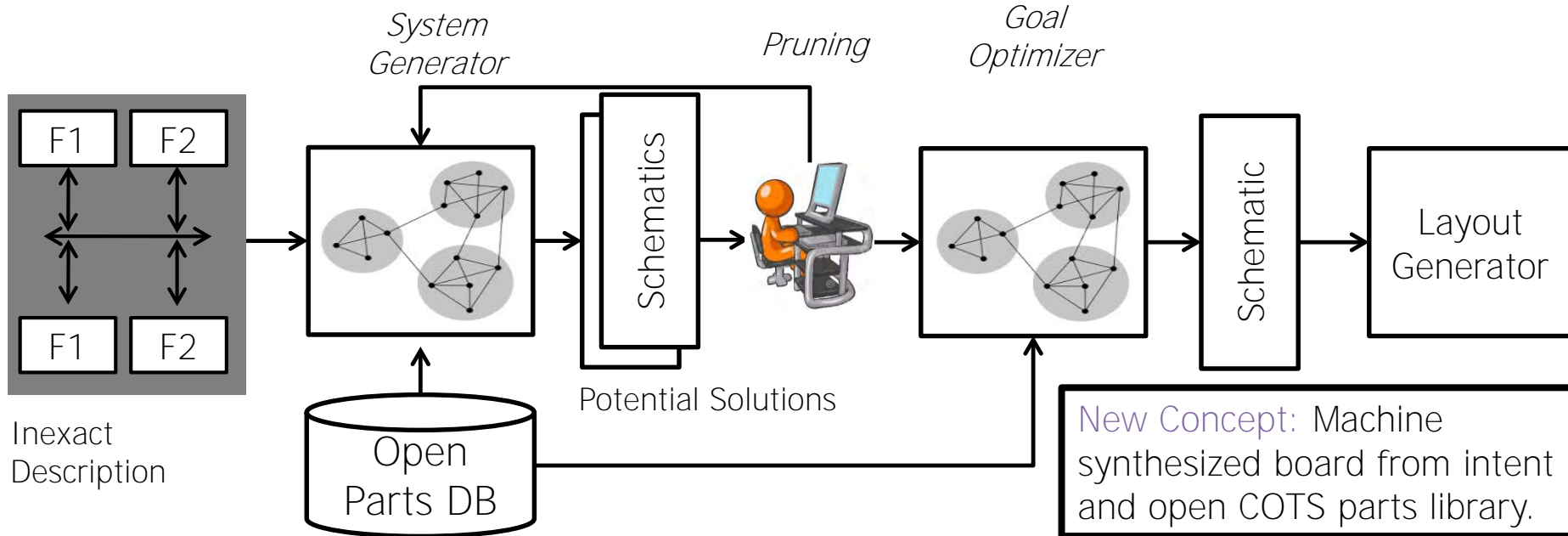


TA2: Reinventing Board Development

Today



IDEA



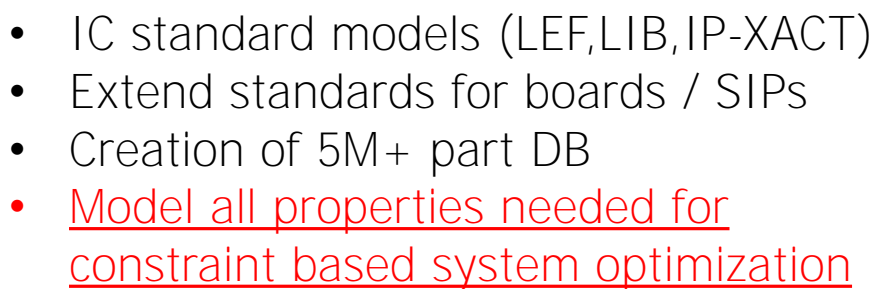


Product Folder Sample & Buy Technical Documents Tools & Software Support & Community Reference Design

TPS65216

TPS65216x Power Management for ARM® Cortex™-A8/A9 SoCs and FPGA

11



- IC standard models (LEF, LIB, IP-XACT)
- Extend standards for boards / SIPs
- Creation of 5M+ part DB
- Model all properties needed for constraint based system optimization



IDEA Design Automation Research Teams

Team	PI	Open Source	Deliverable
Cadence	David White	No	Analog & digital layout generator
Northrop Grumman	Daniel D'Orlando, Jonathan Bachrach	Yes	Board generator
UCSD	Andrew B. Kahng	Yes	Digital circuit layout generator
University of Texas	David Pan	Yes	Analog circuit layout generator
University of Minnesota	Sachin Sapatnekar	Yes	Analog circuit layout generator
University of Utah	Pierre-Emmanuel Gaillardon	Yes	Logic synthesis tool
Purdue University	Dan Jiao	Yes	Parasitic extraction tool
Yale University	Rajit Manohar	Yes	Asynchronous circuit design tool
University of Michigan	David Wentzloff	Yes	System-On-Chip synthesis tool
Princeton University	David Wentzlaff	Yes	IDEA design advisors
University of Illinois	Martin Wong, Tsung-Wei Huang	Yes	Static timing analysis tool



A True Silicon Compiler Will Disrupt the Industry



~~Chip Layout Army~~

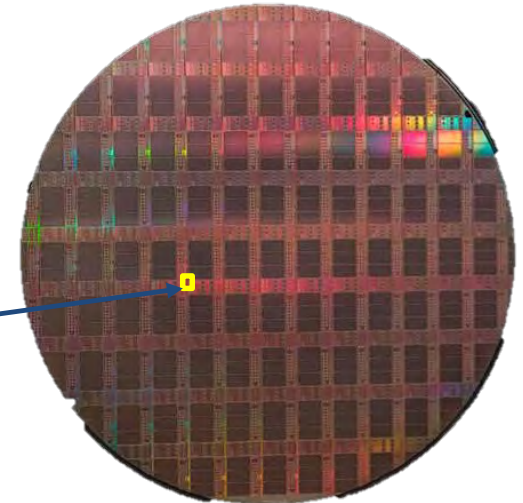
Massive cloud computing

A General Purpose Silicon Compiler:

- Removes expertise barrier to democratize access to silicon technology
- Replace finite human time with machine cycles

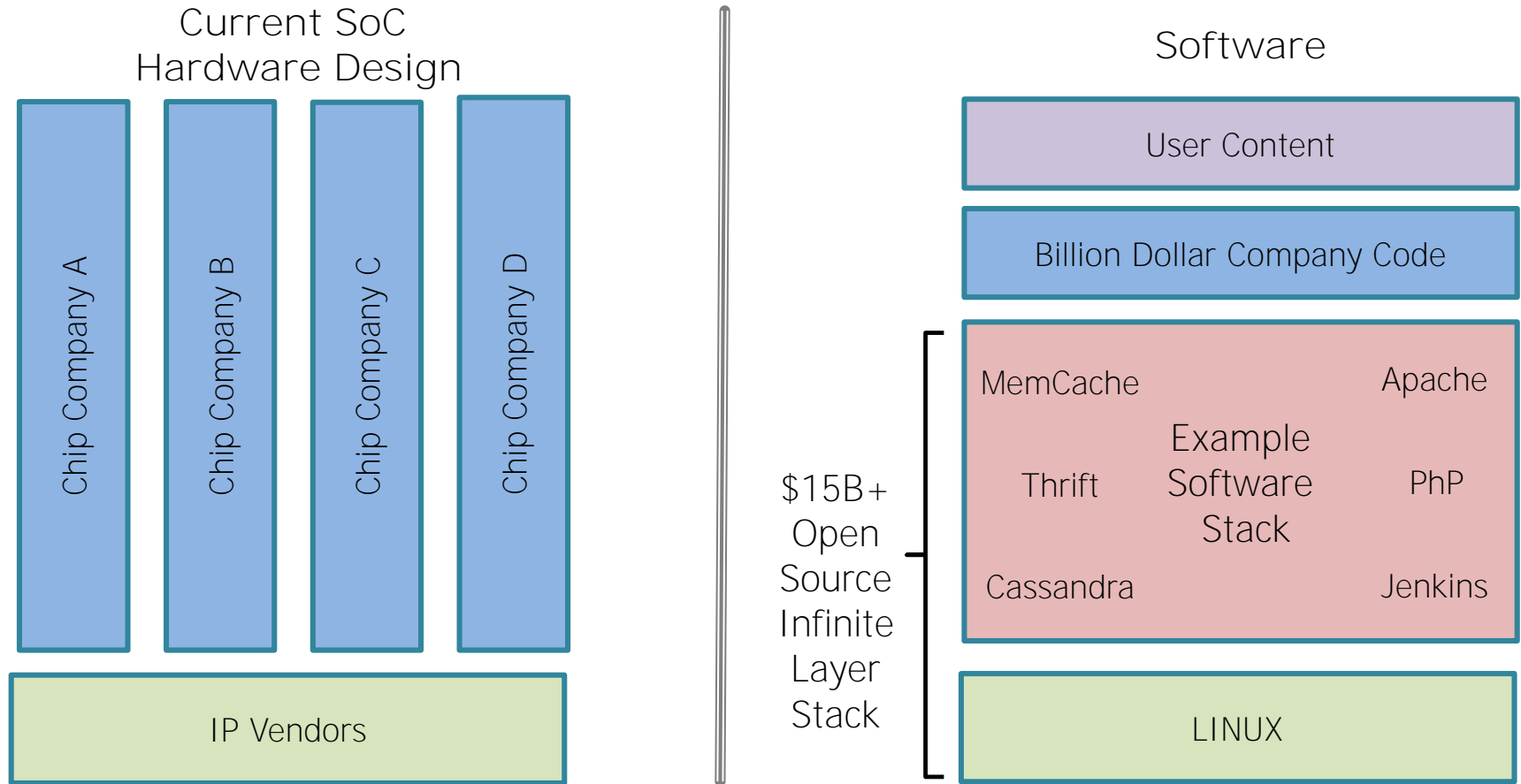
Outcome:

- Makes it practical to specialize for "N=1"
- Reach beyond the horizon, across the chasm,...





POSH: Posh Open Source Hardware



POSH will create a viable open source hardware design and verification ecosystem that enables cost effective design of ultra-complex SoCs.

	Software	Hardware
Programmers	Millions	Thousands
Writing Code	Easy	Hard
Reading Code	Hard	Very hard
Debugging	Hard	Near impossible
Cost of bugs	Low	Very high

What technologies are needed to make open source hardware viable?



POSH program structure



Not funded

Level	Description
L3	Accessible open API hardware emulation and prototyping platforms
L2	Scalable open API mixed accuracy simulation tools
L1	Formal tools for assessing relative and absolute quality of hardware library modules

Increasing levels of assurance

L3: Emulation & Prototypes

L2: Simulation

L1: Formal Analysis



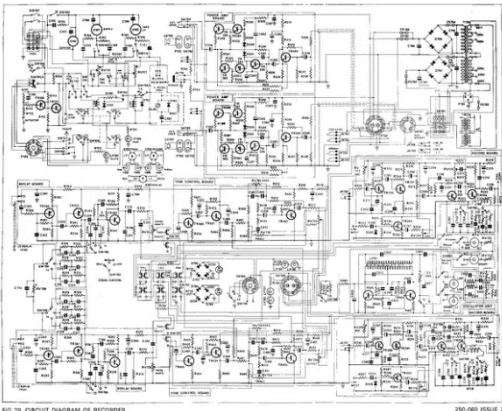
Digital Circuit IP Blocks
FPGA Fabric
Multi-core 64-bit RISC-V processor sub-system
GPU (OpenGL ES 3.0)
PCI Express Controller
Ethernet Controller
Memory Controllers
USB 3.0 Controller
MIPI Camera Serial Interface controller
CPU Subsystem
H264 encoder/decoder
AES256 encrypt/decrypt
SHA-2/SHA-3 accelerator
Secure Digital Controller
High Definition Multimedia Interface
Serial ATA Controller
JESD204B Controller
NAND Flash Controller
CAN Controller

Mixed Signal Circuit IP Blocks	Description
Standard I/O interfaces PHYs	DDR, PCIe, SATA, USB, XAUI, CPRI
PLL	Range: 10MHz – 10GHz
DLL	Range: 10Mhz – 10GHz
Analog to Digital Converters	Range: 1 – 10,000 MSPS
Digital to Analog Converters	Range: 1 – 10,000 MSPS
Voltage Regulators	Input: 1.8V – 12V, Output 0.25V – 1.8V
Monitor circuits	Temperature, voltage, process

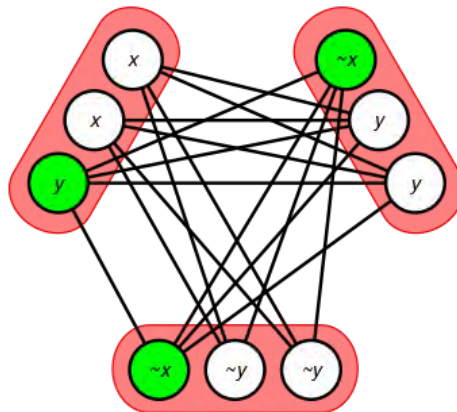
How can we cost effectively develop and maintain a high quality catalog of portable open source digital and analog components?



POSH: Expected Program Results



Open source analog IP



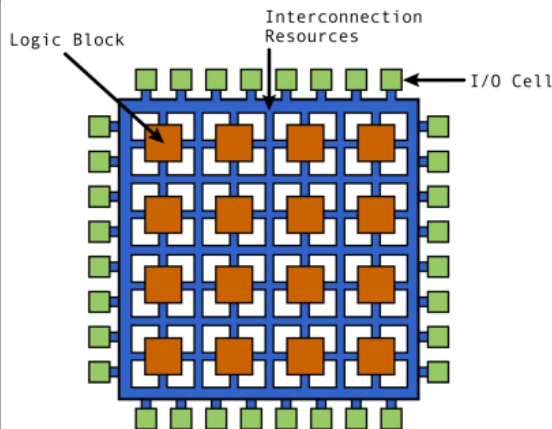
Signoff level validation



Commercial open source design community



Multicore RISC-V



Open source FPGA Chips



"Linux for SoC Design"



POSH Open Source Research Teams

Team	PI	Open Source	Deliverable
Xilinx	Edgar Iglesias	Yes	Hardware/Software co-simulation
Synopsys	Oleg Raikhman	Yes/No	Mixed Signal Emulation
Sandia	Eric Keiter	Yes	Parallel Mixed Signal Simulation Tool
Stanford University	Clark Barrett	Yes	Formal Verification Tools
University of Washington	Richard Shi	Yes	PDK independent analog design
University of Washington	Michael Taylor	Yes	High performance RISC-V processor
USC	Tony Levi	Yes	High performance analog circuits
University of Utah	Pierre-Emmanuel Gaillardon	Yes	FPGA Generation Tool
LeWiz	Chinh Le	Yes	Ethernet Controller
Princeton University	David Wentzlaff	Yes	Open FPGA fabric
Brown University	Sherief Reda	Yes	SoC PVT sensor circuits



- Design a state of the art open source mixed signal ~~100mm²~~ SoC
- Design an interesting AND important capability!
- Perform early testing of TA1 and TA2 technology
- Integrate majority of TA2 IP blocks in design
- Design any specialized components required for SoC
- ~~• Demonstrate a 10X+ SWAP-C improvement over existing COTS solutions~~
- DARPA will provide fabrication support through MPWs

Can we design a provably secure 100% open source System-On-Chip?

```
$ git clone https://github.com/darpa/idea
$ git clone https://github.com/darpa/posh
$ cd posh
$ make soc42
```



24 Hours



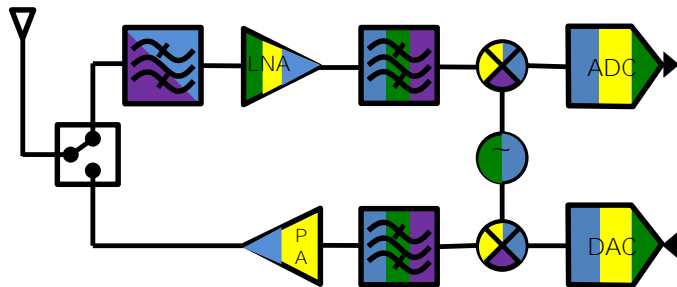
www.darpa.mil

Common Heterogeneous Integration and IP Reuse Strategies (CHIPS)

Andreas Olofsson
Program Manager, DARPA/MTO

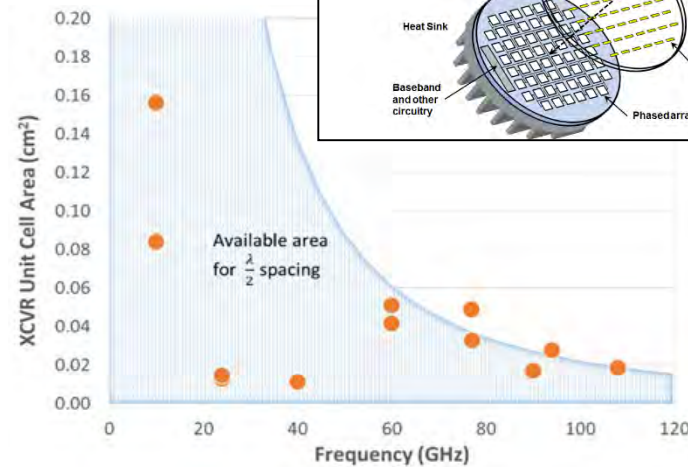
Arlington, VA
Dec 19, 2018



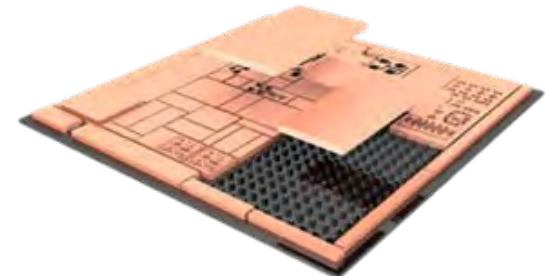
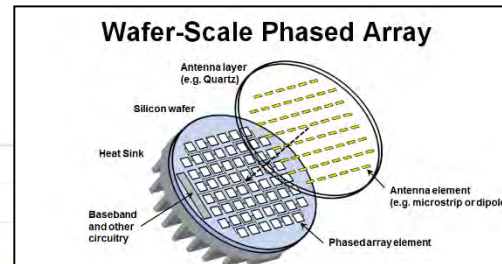


- Si CMOS/SiGe BICMOS
- InP HBTs/HEMTs
- GaN HEMTs
- RF MEMS/High-Q passives

Best of breed technology

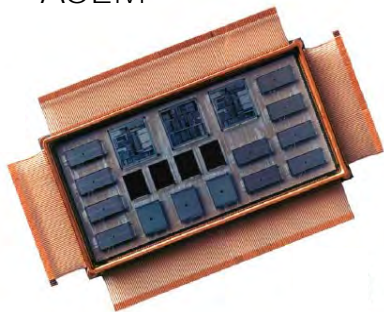


Density of integration

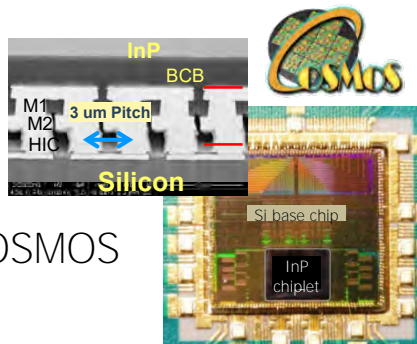


Modular design

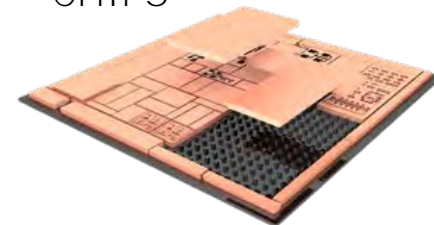
ASEM



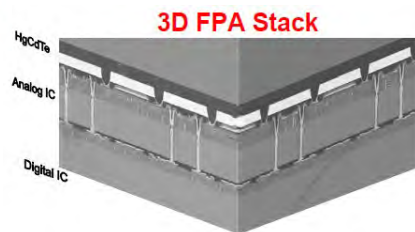
COSMOS



CHIPS



VISA



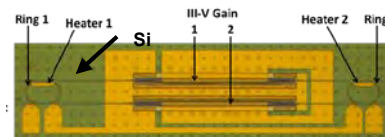
3D-IC



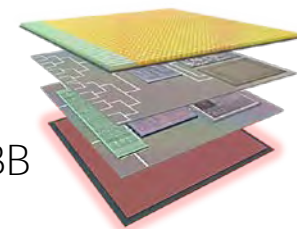
DAHI



E-PHI



MOABB



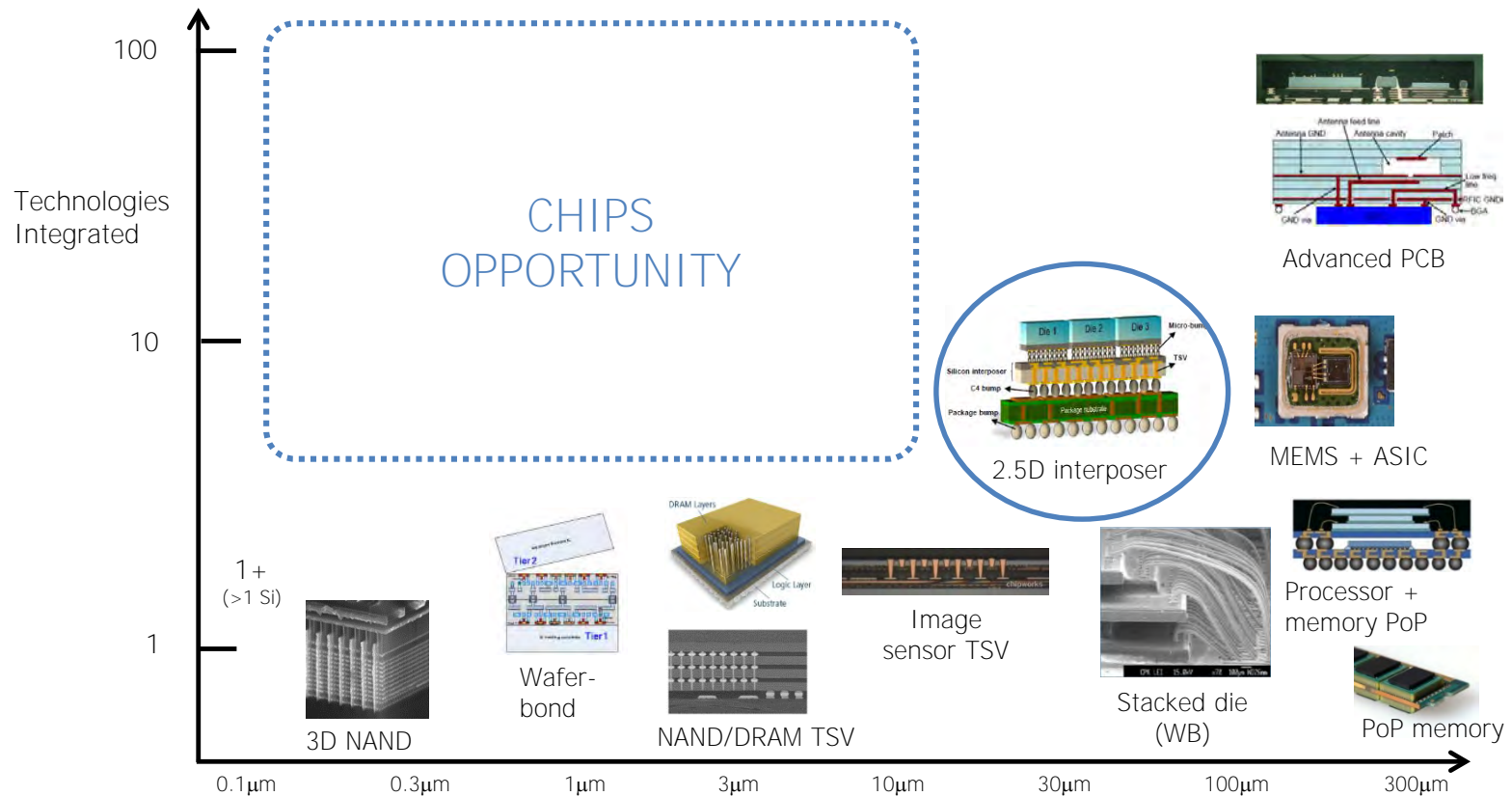
ASEM: Application Specific Electronic Modules
 E-PHI: Electronic-Photonic Heterogeneous Integration
 VISA: Vertically Integrated Sensor Arrays
 COSMOS: Compound Semiconductor Materials on Silicon
 DAHI: Diverse Accessible Heterogeneous Integration
 MOABB: Modular Optical Aperture Building Blocks
 CHIPS: Common Heterogeneous Integration and IP Reuse Strategies

1990s

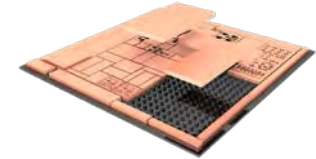
2000s

2010s

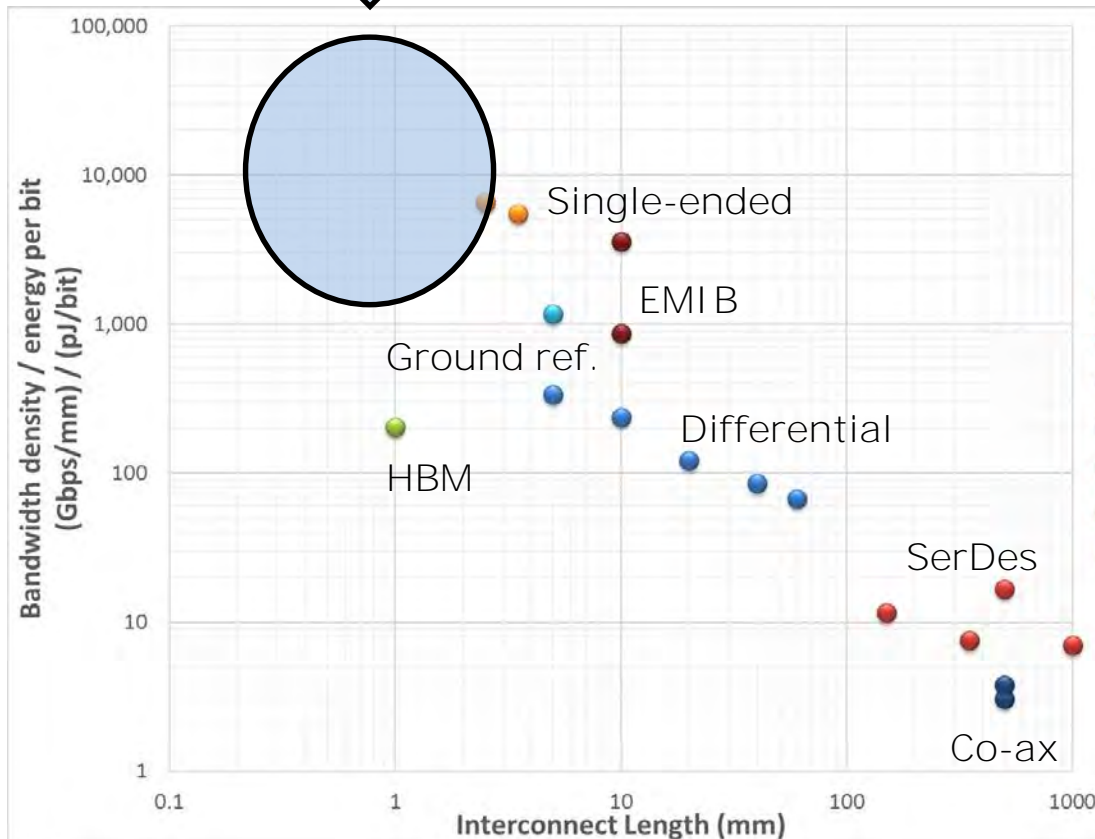
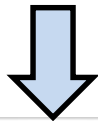
2020s



Sources: Samsung, Tezzaron, Micron, Hynix, Sony, TSMC, Prismark, STATS ChipPAC, Georgia Tech, IEEE, STMicro, Finetech



CHIPS Target



CHIPS Program Interface Standard Metrics

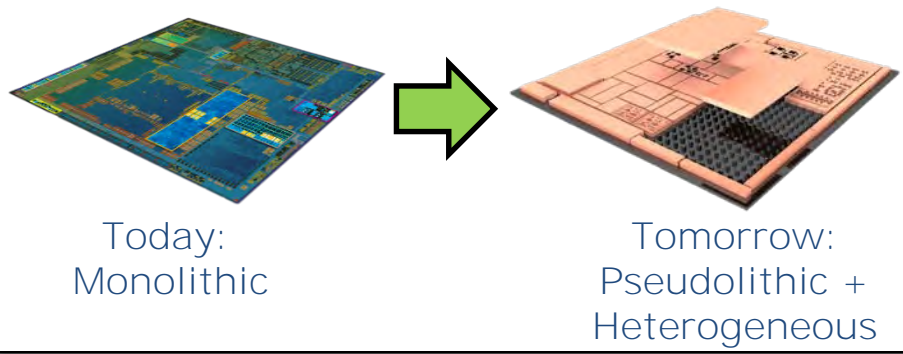
Data rate	10 Gbps
Energy efficiency	< 1 pJ/bit
Latency	< 5 ns
Bandwidth density	> 1000 Gbps/mm

Sources:

1. 2016 JSSC, Dehlaghi
2. 2013 JSSC, Poulton
3. 2012 JSSC, Dickson
4. 2013 JSSC, Mansuri
5. 2016 ECTC, Mahajan

CHIPS interface is one of many possible routes for efficient interdie communications

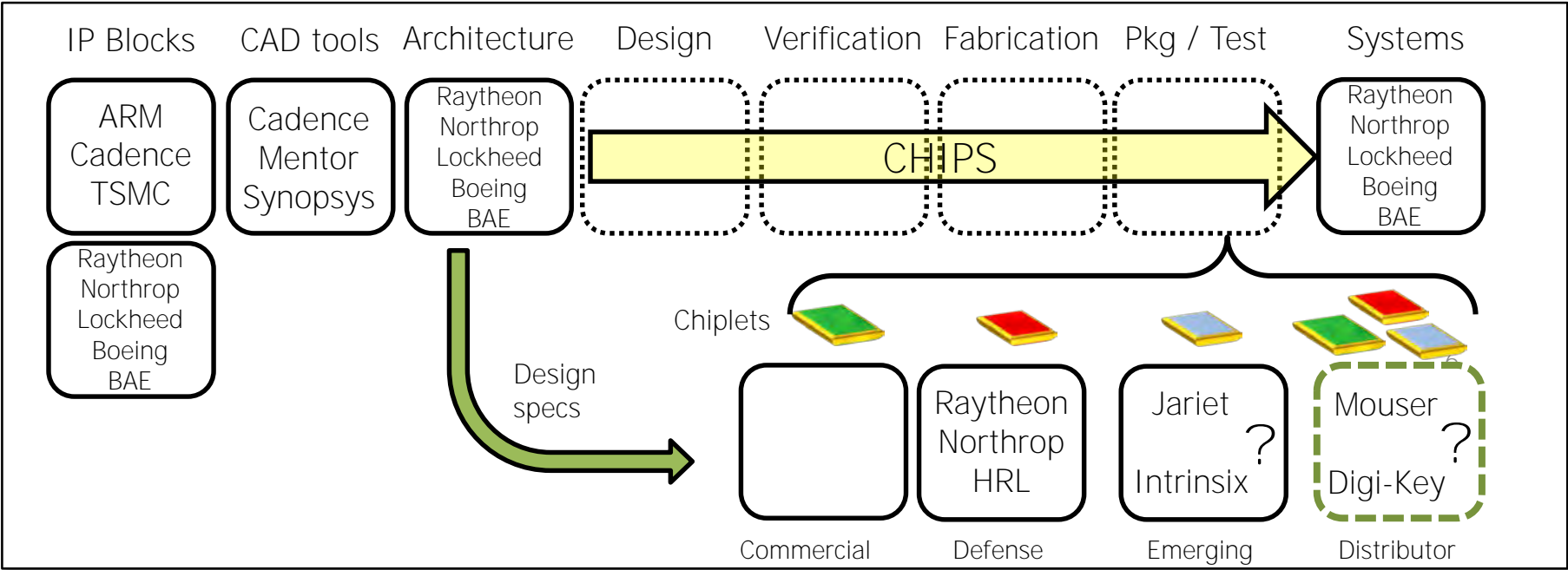
DESIGN

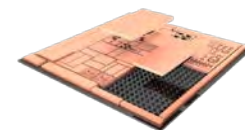


TECHNOLOGY

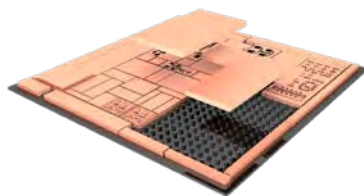
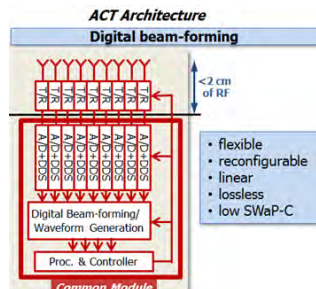
- ✓ A universal CHIPS interface standard
- ✓ SOTA manufacturing for DoD
- ✓ A critical set of IP chiplets

CHIPS END STATE





TA1: Modular Digital Systems

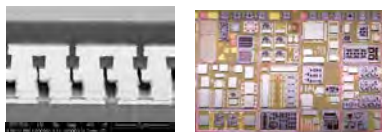


TA3: Supporting Technologies

Chipletized IP

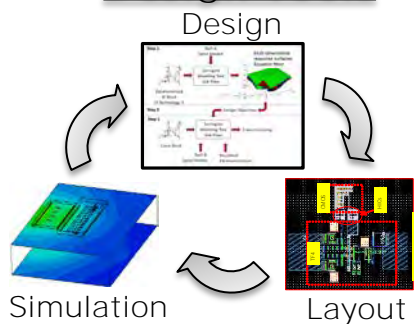
ISP	SerDes	Controller
Audio signal	USB	DRAM
Digital signal	PCIe	SRAM
Compression	GPU	Flash
GPU	CPU	MP

Assembly Methods



Fine pitch Heterogeneous interconnects integration

Design Tools

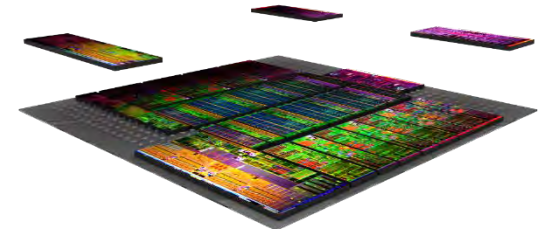


	Team	Focus
TA1	Boeing	NOC, 10um, system
	Intel	FPGA Platform, Standard
	LM	Obsolescence
	Michigan	Deep learning chiplet
	NGMS	ACT chip disaggregation
TA3	Cadence	2.5D modeling
	GTech	NOC, PMIC, PNR
	Intrinsix	Root of trust chiplet
	Jariet	Quad 64 Gsps ADC/DAC
	Micron	GDDR6 chiplet
	NCSU	RISC-V SoC chipletizing
	NGAS	Standard I/O chiplets
	Ferric	Interposer PMIC
	ADI	10 Gsps ADC
	UCLA	10um pitch Die on Wafer

1. Address the System-On-Chip dilemma.
Enable system companies with limited chip design expertise to leverage the power of advanced CMOS manufacturing

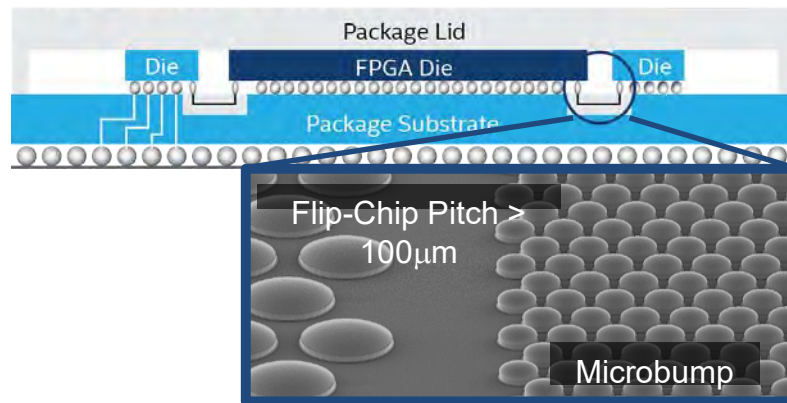


3. Extend Moore's law
Scale out and scale down while managing yield

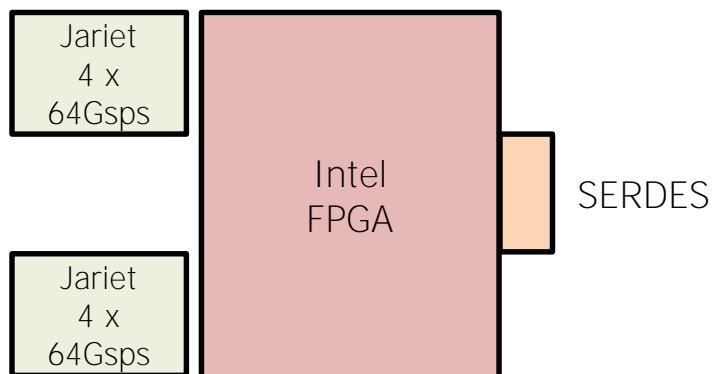


Intel production proven manufacturing

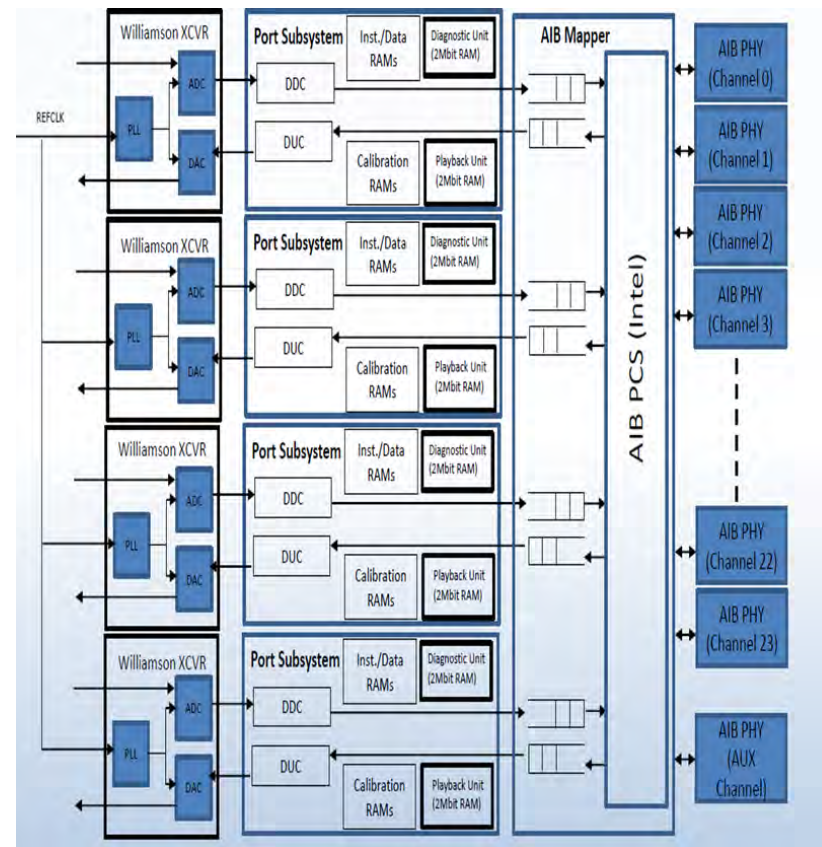
Intel® Stratix® 10 FPGAs and SoCs with Intel EMIB

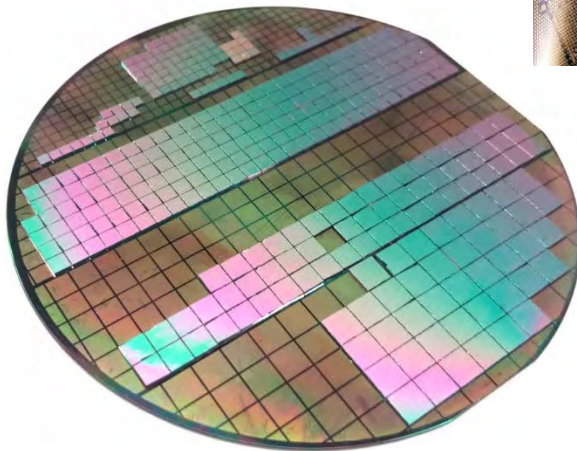
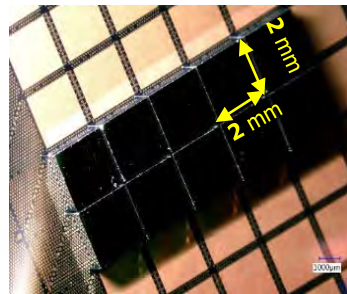
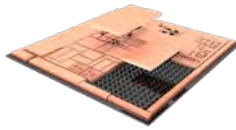


Intel/CHIPS MCM using EMIB Technology with AIB interface standard



Jariet direct RF sampling at up to 64Gsps, with quad channel 10-bit ADC/DAC IP (existing, lab-proven ACT IP is being reused on CHIPS)

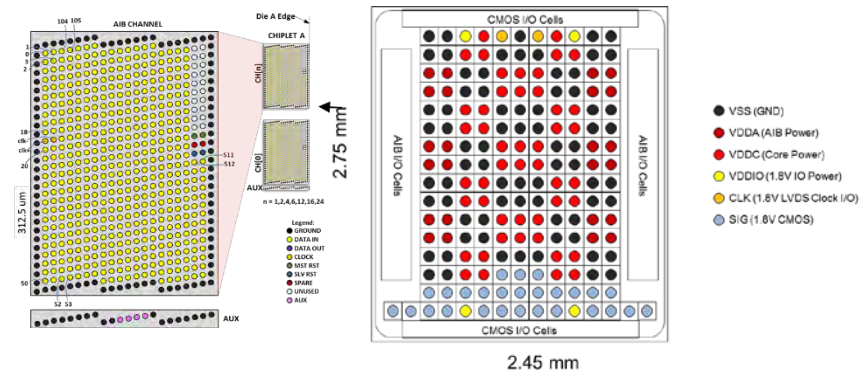
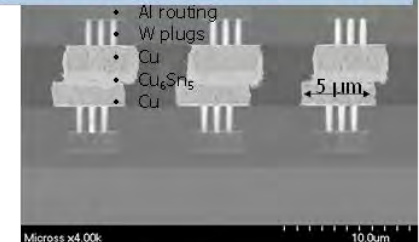
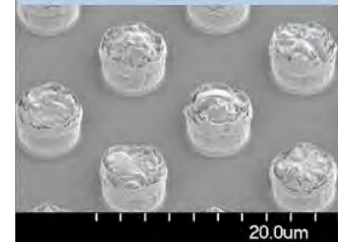




UCLA:

- Si IF fabricated Dual Damascene process
- ~370+ dielets assembled (4mm² - 25mm²)
- 10μm pitch (±1 μm alignment; θ < 6m deg)
- 100μm spacing
- >3000mm² total dielet area
- Passivated with Parylene C
- Close collaboration with Kulicke & Soffa

Initial Cu Feasibility Demonstrated June 2018

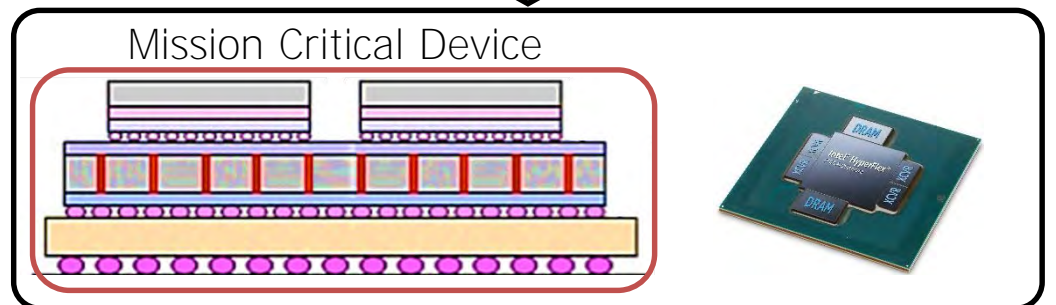
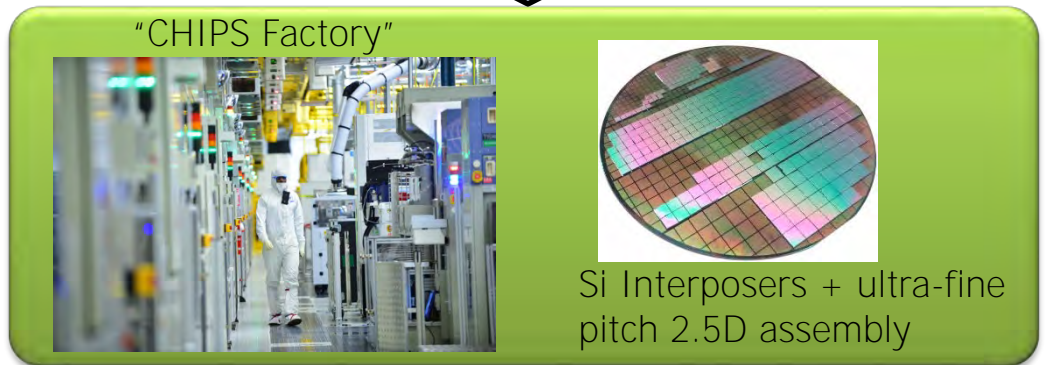
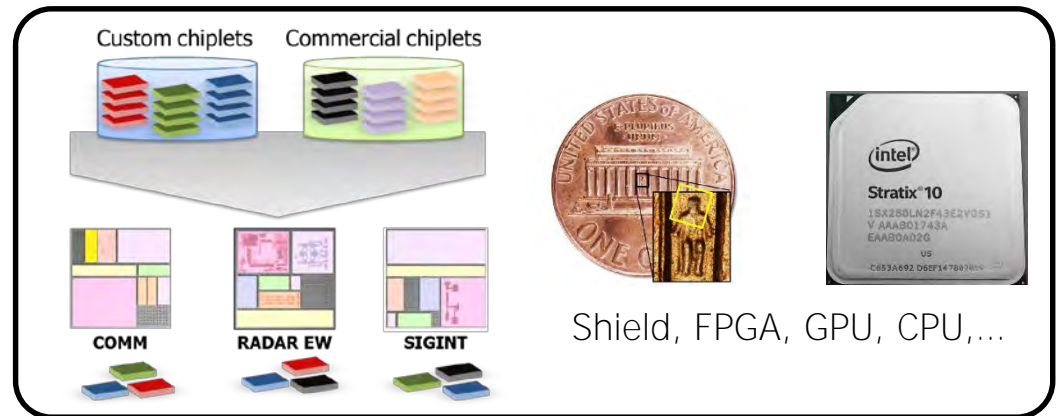


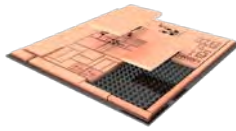
- Northrop Grumman & Micross demonstrated ultra-fine pitch interconnect required for high-speed, highly parallel interface
- CHIPS is developing options for DoD-scale manufacturing via MPWs, foundry-agnostic processes, die-level processing, domestic interposer sources



CHIPS is a once in a generation disruption to SWAP limited electronics:

- Better than monolithic performance through heterogeneous integration
- PCB-like cost & time scales

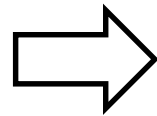




10^{12} bits/sec
 10^{-13} Joule/bit

CHIPS
 Interface

Plug and Play
 Standard



ADC

CHIPS

Front
 End
 DSP

CHIPS

Balanced
 Performance &
 Flexibility

FPGA

CHIPS

Low Latency
 ML
 Accelerator

$\ll 1$ ms
 Latency?



www.darpa.mil



Distribution Statement "A" (Approved for Public Release; Distribution Unlimited)

Foundations Required for Novel Compute (FRANC)

Young-Kai Chen, PM

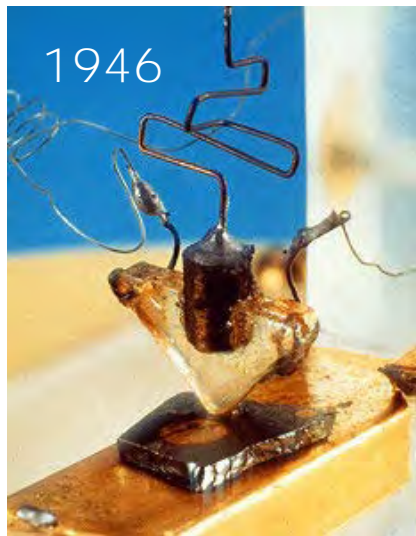
Electronics Resurgence Initiative: Defense Applications

December 19, 2018

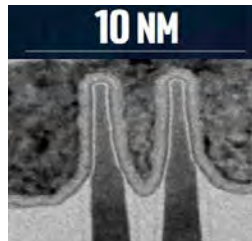


A robust von Neumann architecture powered by the CMOS scaling with Moore's Law

Highly Scaled CMOS

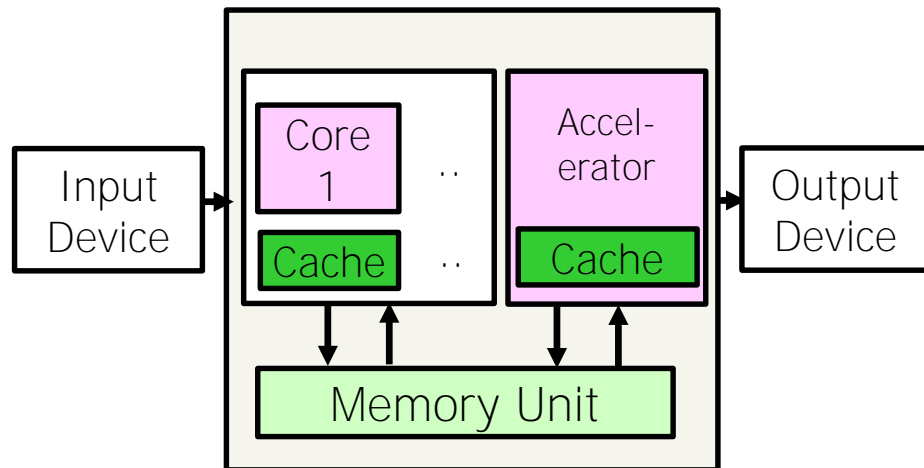


1946



10 NM

Parallelization to reduce latency

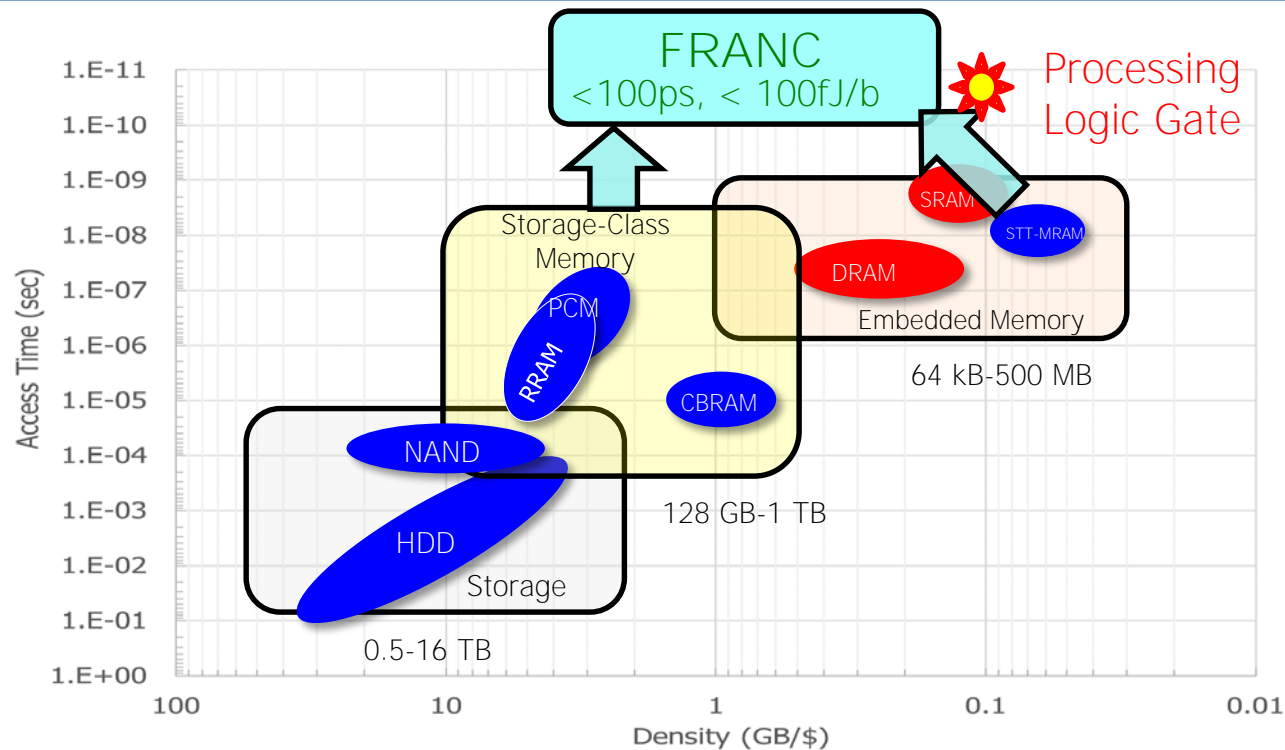


Distribution Statement "A" (Approved for Public Release, Distribution Unlimited)



Today's processor speed is 100x faster than memory

FRANC utilizes new materials and devices to make 10x advances in embedded non-volatile memories with speed as SRAM and density as storage-class memory



- Non-Volatile
- Volatile

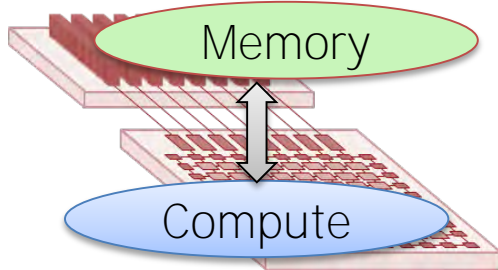
Reference: Siva Sivaram, Western digital Corporation, August 9, 2016, Flash Memory Summit

Distribution Statement "A" (Approved for Public Release, Distribution Unlimited)



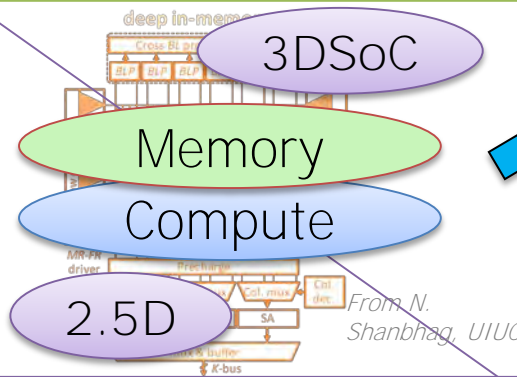
FRANC enhances memory-centric computing architecture

Present von Neumann Computing



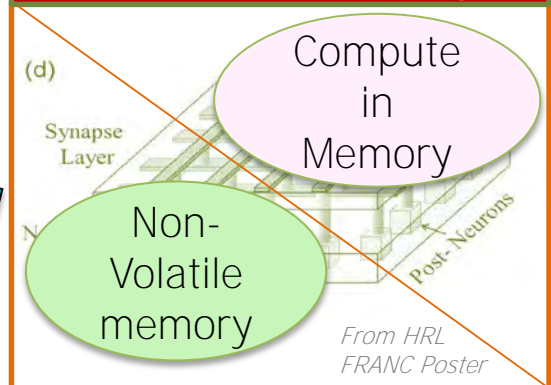
From S. Mitra at Stanford

Enhanced von Neumann Computing

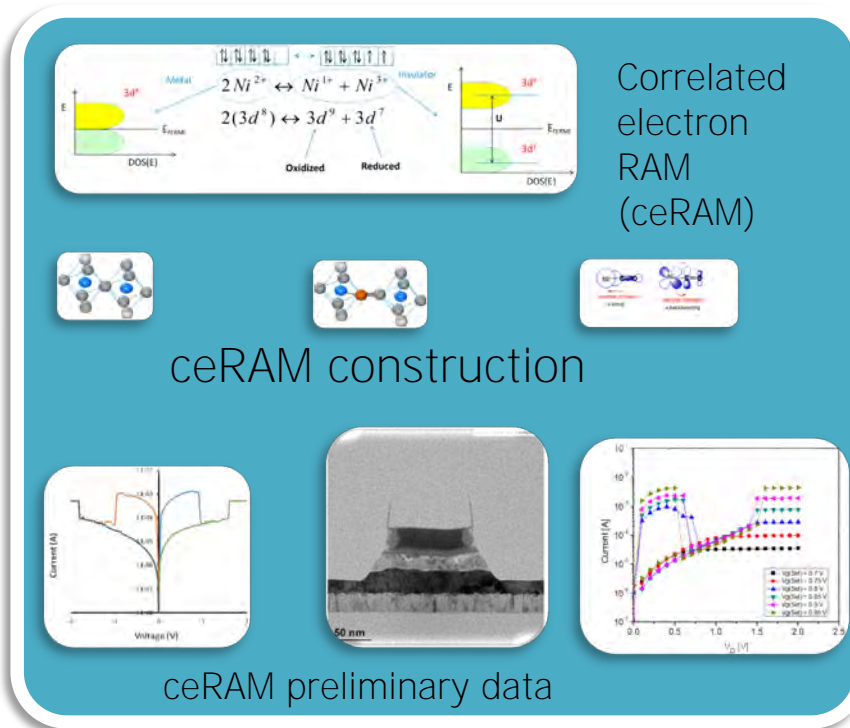


From N. Shanbhag, UIUC

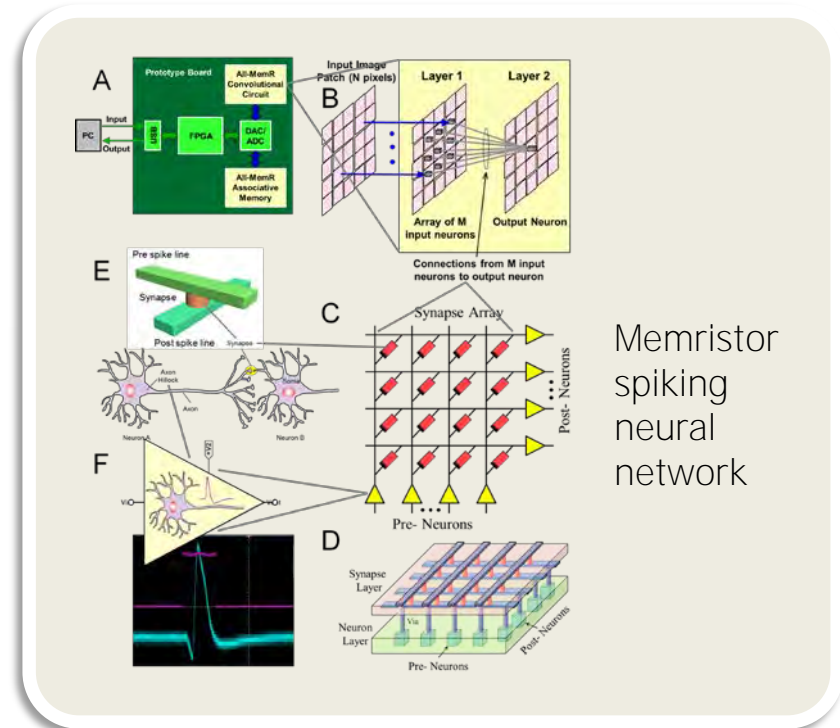
Emerging Memory-Centric Computing



*From HRL
FRANC Poster*



Applied Materials



HRL Laboratories



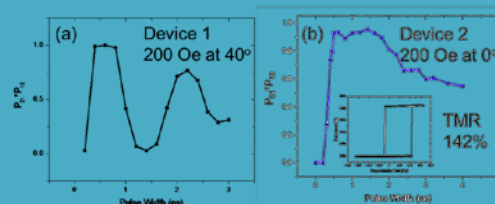
Voltage-Controlled Magnetic tunnel junction memory

- Ultra fast switch
- Low switch energy
- Small feature size



$K_{eff}(V) < K_{eff}(0)$
 $K_{eff}(V) < K_{eff}(0)$
 Damping ~ 0.002
 TMR 400%-500%
 Switching mechanism: STT + Cross-VDMA

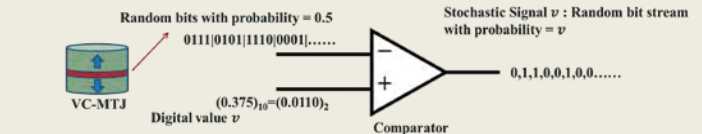
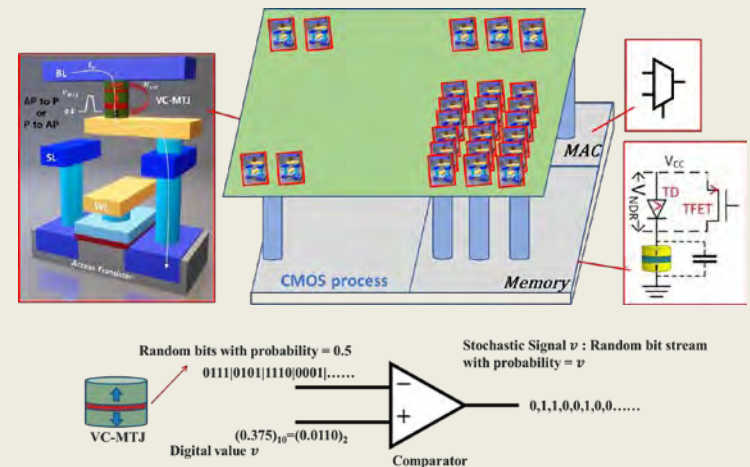
10 nm p-MTJ,
Energy-Delay =
100 aJ · 10 ps



Preliminary data

University of Minnesota

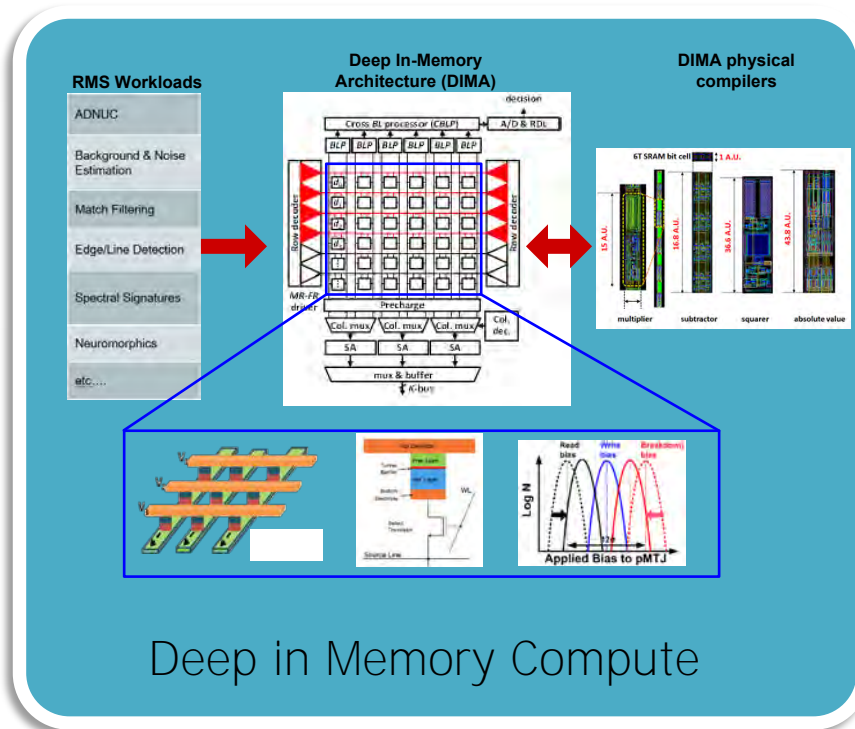
Magnetic tunnel junction memory for stochastic computing



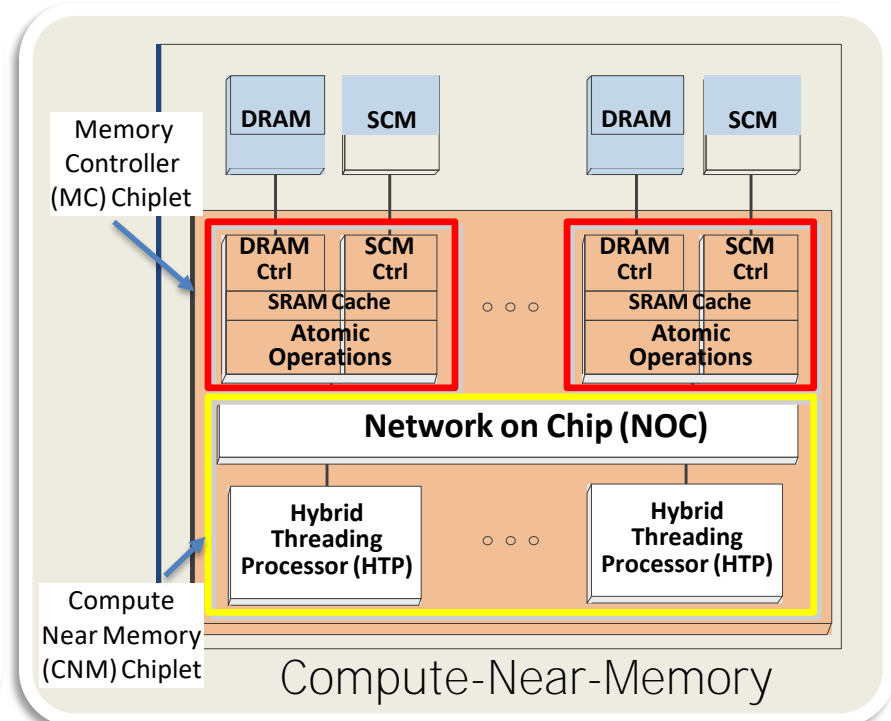
University of CA, Los Angeles



Novel Computing Architectures



University of Illinois at UC



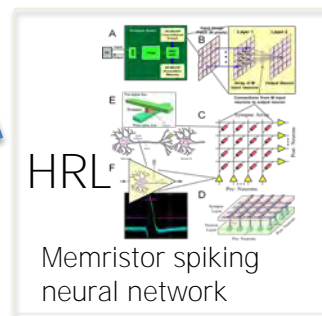
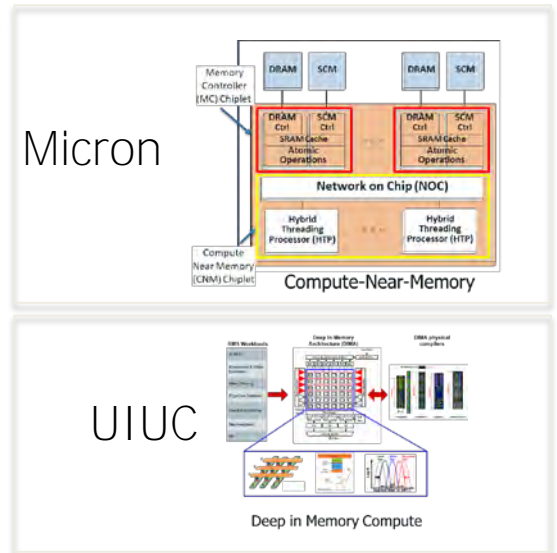
Micron



Areas of interest for transition

Leverage advantages developed under FRANC:

- In and near memory processing
 - Image/signal processing
 - Radar backprojection
 - Graph processing scalable to large data sets
- Low-SWAP applications using neural network classification

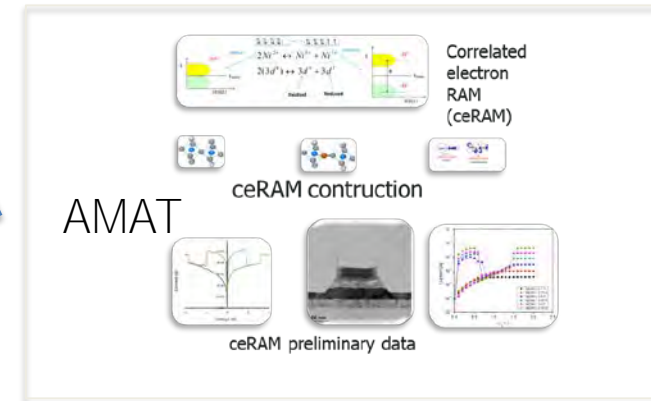
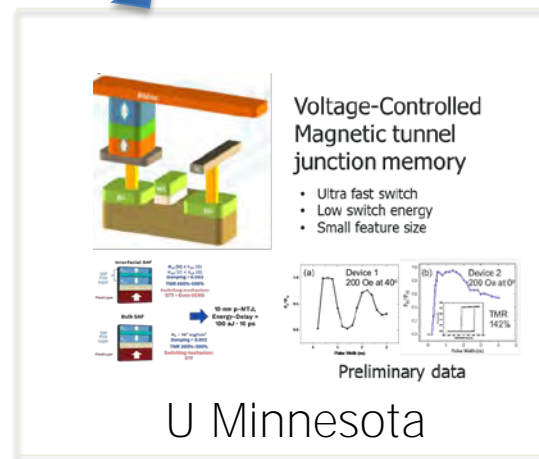
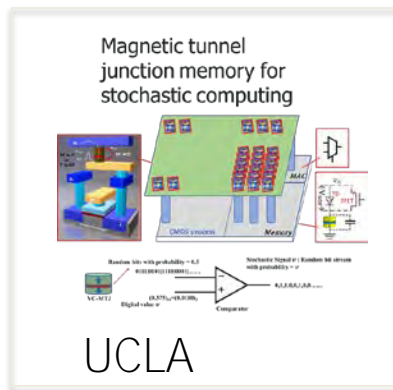




Areas of interest for transition (cont.)

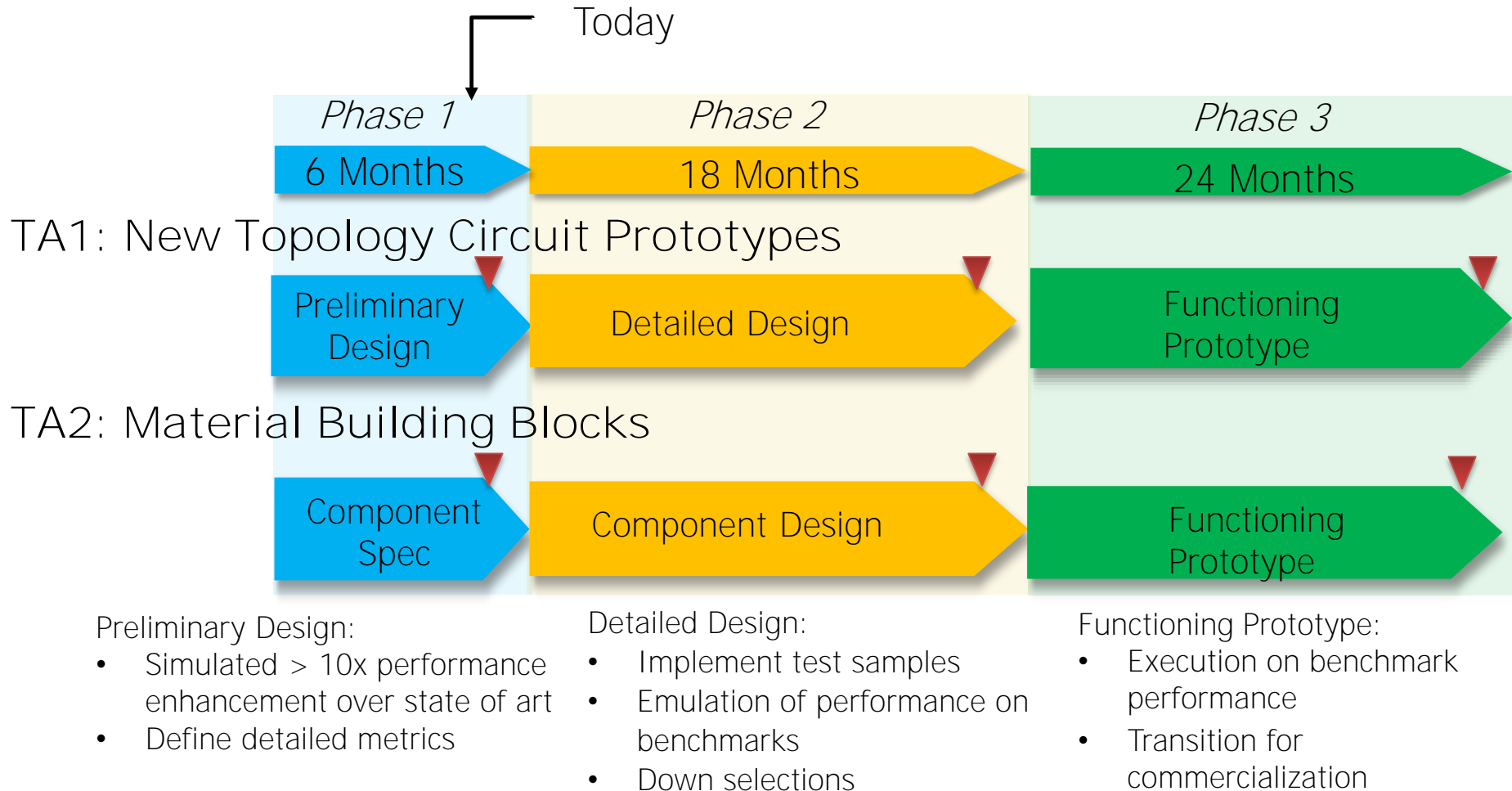
Leverage advantages developed under FRANC (cont.):

- Applications requiring non-volatile memory
 - Multi-state logic
 - Voltage-Controlled MRAM
- Fault-tolerant and stochastic computing





FRANC program structure and metrics





Submissions

Guidelines for FRANC transitions:

- Identify FRANC partner and advantage of technology
- Illustrate problem and clear benefit to using FRANC technology
- Show the DoD-relevant application and what is the advantage over existing technology
- Abstracts are encouraged

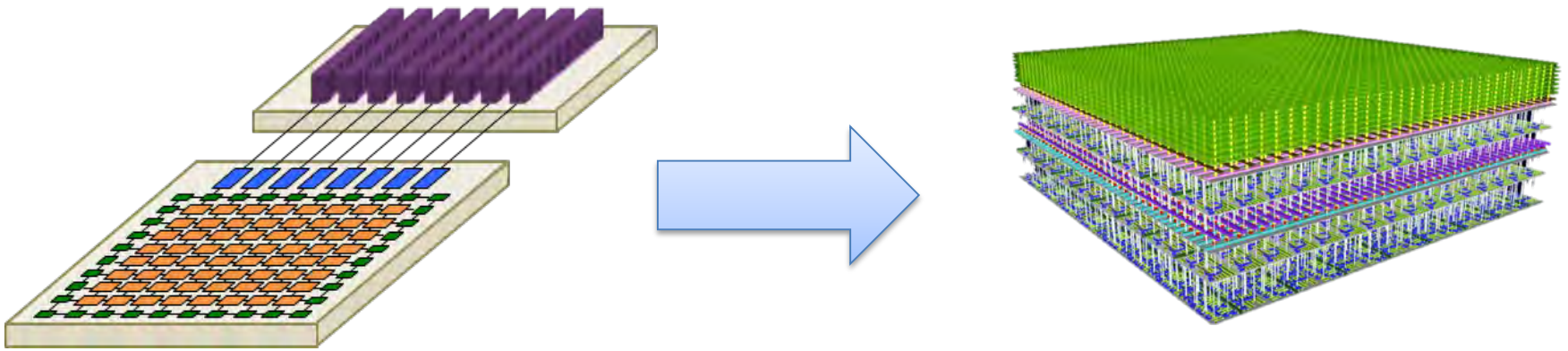
Contacts for PIs:

Institution	Name
HRL	Dr. Wei Yi
Applied Materials	Dr. David M. Thompson
Micron	Mr. Glen Edwards
UCLA	Prof. Sudhakar Pamarti
UMN	Prof. Jian-Ping Wang
UIUC	Prof. Naresh Shanbhag



www.darpa.mil

3 Dimensional Monolithic System on a Chip (3DSoC)

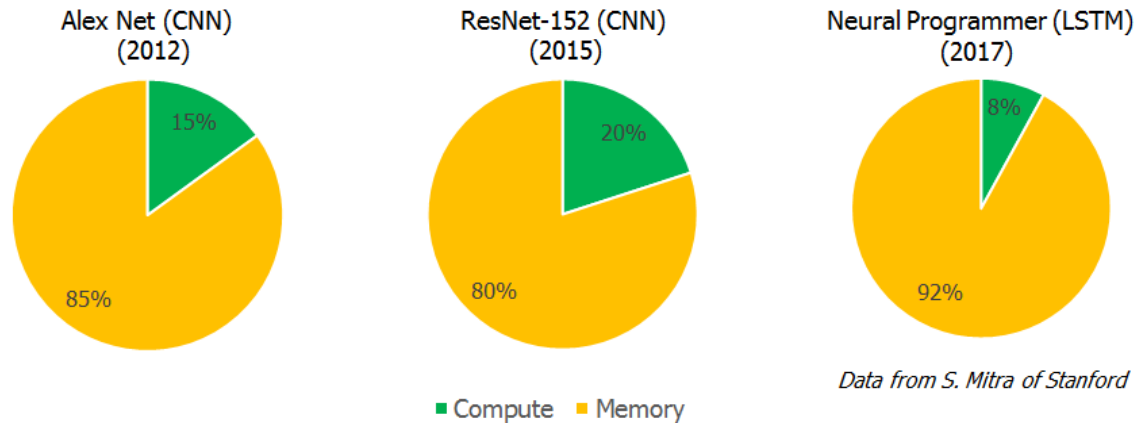


Develop novel monolithic 3D fabrication technologies that enable new architectures to drive a >50X improvement in SoC performance at power



December 19, 2018

Data from S. Mitra of Stanford for a 7nm instantiation of state-of-the-art Machine Learning accelerators



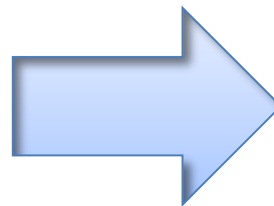
Note: These data representative of systems that benefit from massive caching, parallelism, and pipelining

Amdahl's Law

$$\text{Overall Speedup} = \frac{1}{(1-F) + \frac{F}{s}}$$

F = Fraction enhanced

s = Speedup of enhanced fraction



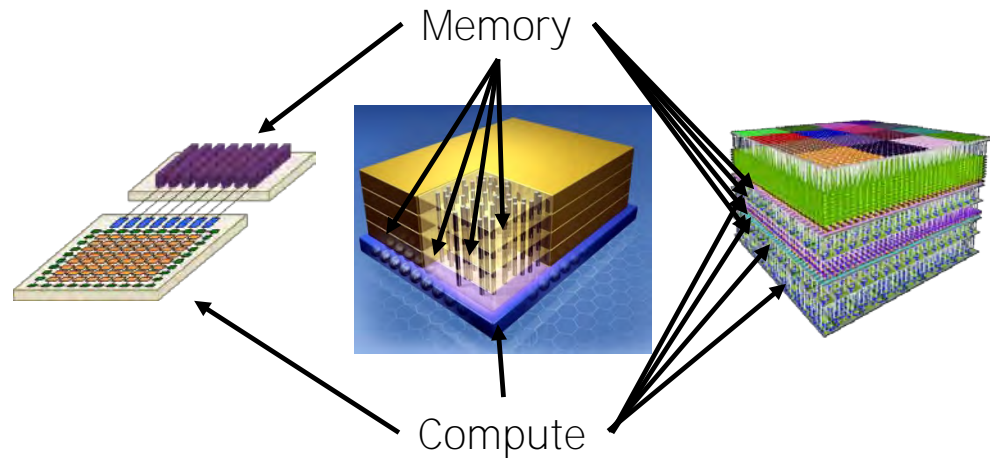
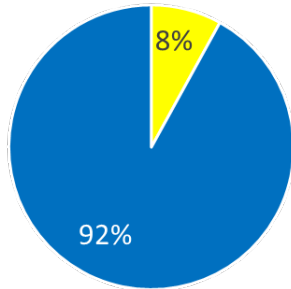
Compute Speedup is throttled by memory since memory access is not speeding up

Memory Speedup	Compute Speedup	Fraction Compute	System Speedup
1X	100X	8%	9%
10X	1X	8%	580%

Arbitrary increases in speedup of computation have limited impact on system performance unless the memory bottleneck is addressed



Addressing the Memory Limitation



Production
2D

Development
3D TSV Package

DARPA
3DSoC

Memory Access Parameter	2D		3D TSV Package		3DSoC
Total I/O	512		1K		33K
Max Bandwidth (Gb/s)	400		1K		46K
Memory access energy (pJ/bit)	52		32		1.5
VDD (Volts System)	1.6		1.2		0.6

3DSoC increases the IO count and bandwidth by >50X from current 2D fabrication architectures



Simulation Results for Machine Learning

LSTM Network	Model Size	Training/Inference	2D at 7nm	
			Benefit 3DSoC at 7nm	Benefit 3DSoC at 90nm
Language Model	2.5 Gbytes	Training	645X	75X
		Inference	626X	73X
Neural Programmer	1 Gbyte	Training	359X	40X
		Inference	493X	55X
Image Captioning	150MByte	Training	367X	41X
		Inference	323X	35X

from S. Mitra of Stanford University

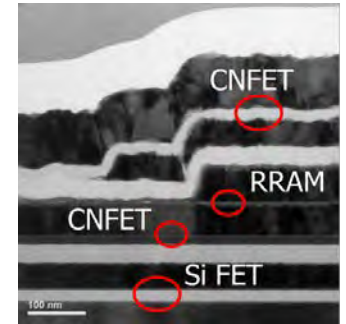
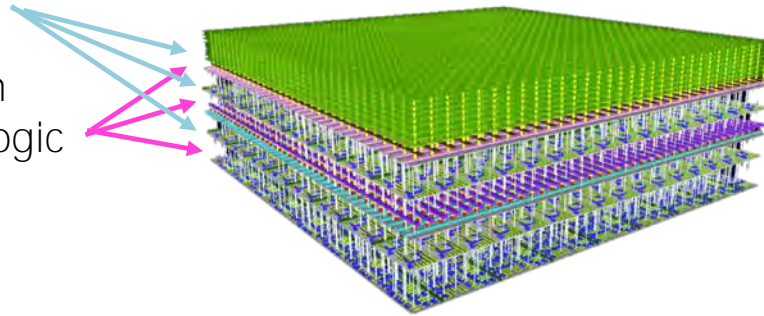
- 2D vs 3DSoC comparison
 - 2D: 7nm technology for accelerator and 4GB of off-chip DRAM main memory
 - 3DSoC: 90nm Carbon Nanotube FET (CNFET) for accelerator and 4GB of on-chip ReRAM (non-volatile) memory
- Uses published traces from an accelerator SoC and the LSTM algorithm
- $\text{Benefit} = (E \cdot t)_{2D} / (E \cdot t)_{3D}^{**}$
- Benefits would be enhanced if design targeted at 3DSoC technology



An Integrated, Monolithic SoC (3DSoC) Solution

An integrated flow that fabricates 3D logic and memory on a single die

12 layers of ReRAM
interspersed with
5 layers of CNFET logic



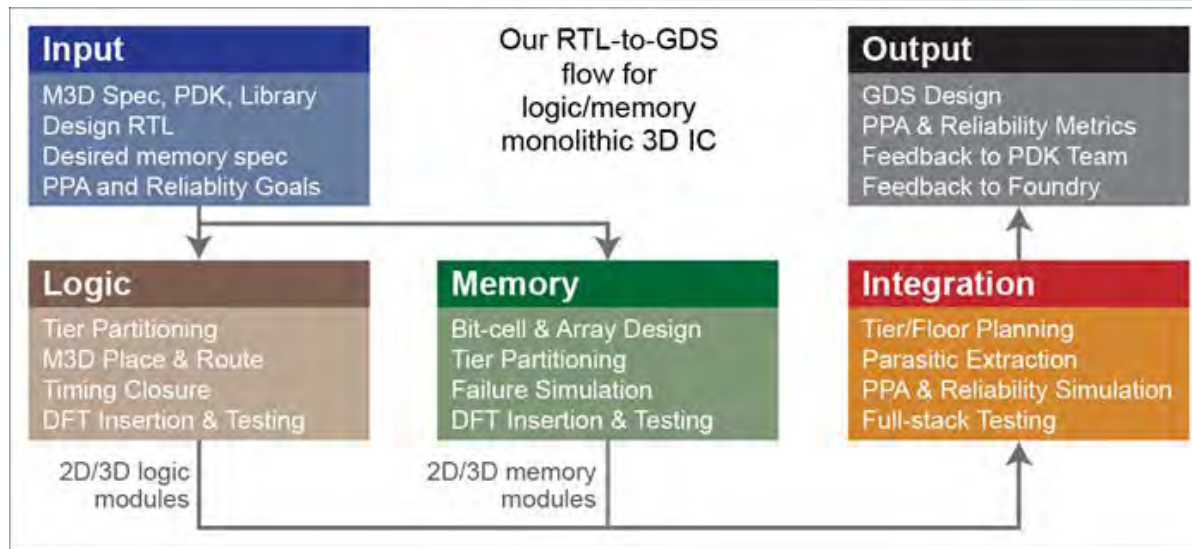
from S. Mitra of Stanford University

Critical characteristics for a monolithic solution

- Must permit new architectures that leverage fast, configurable access to non-volatile main memory
- Stackable 3D logic and memory functions that allow new architectures
 - Low temperature formation
 - Logic AND memory
 - High density of memory – at least 4GB (Giga-Byte)/die
- Possible to fabricate in existing domestic, commercial, high-yielding infrastructure
 - 90nm on 200mm wafers
 - High yield on large SoCs



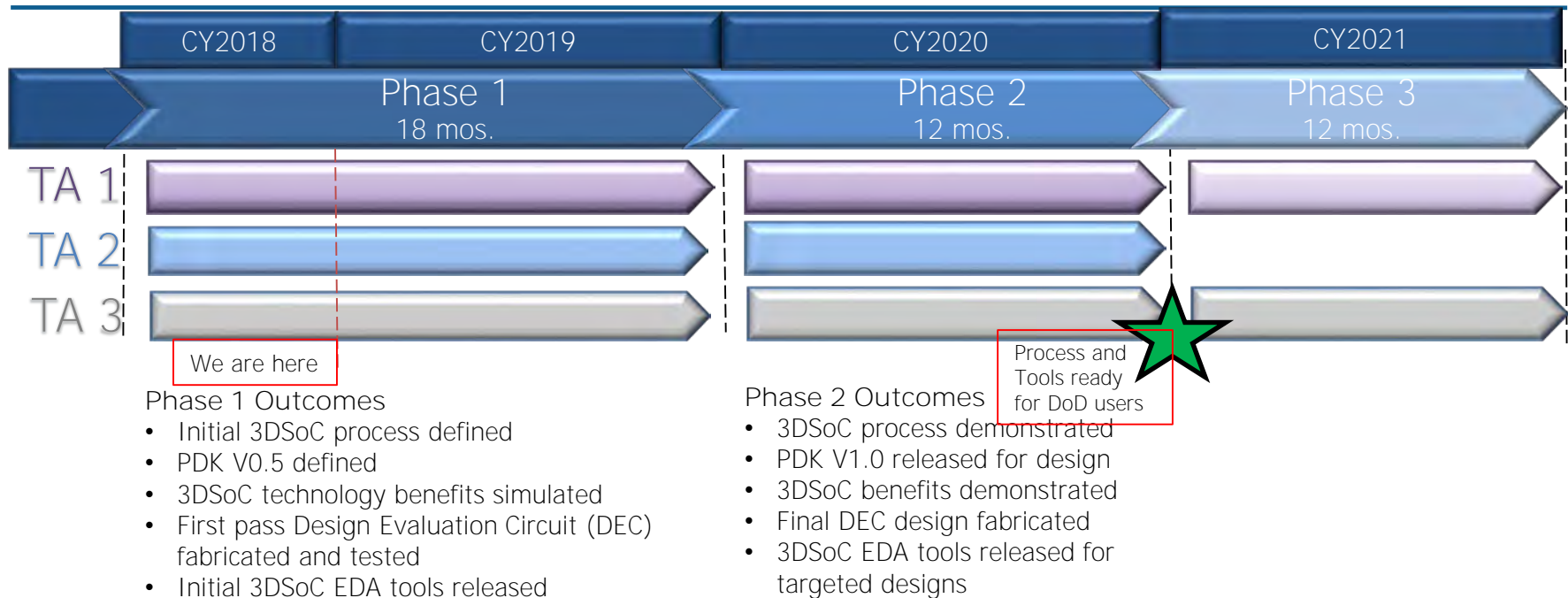
GaTech 3DSoc Design Software Development



- Partitions memory and logic into tiers based on design/technology defined characteristics
- Partitions can be either large or small and can be interspersed
- Augments existing 2D EDA tools as required
- Reliability
- Design for Test (DFT)



3DSoc Program Schedule

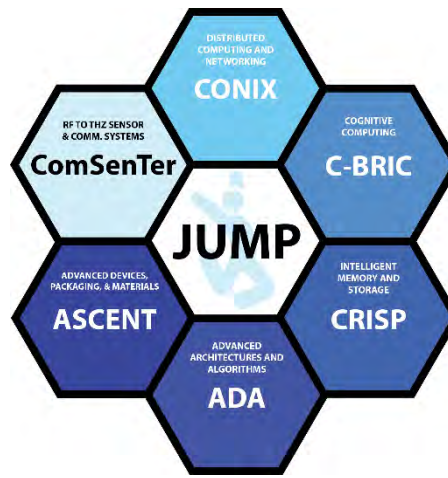


- TA-1: Developing the 3DSoc fabrication process
 - Establish unit processes and flow integration
 - Define the 3DSoc technology PDK
- TA-2: Designing and Implementing the DEC
 - Design 1st and 2nd pass DEC design
 - Foster use of the DEC to drive development and yield
- TA-3: Developing the 3DSoc EDA design flow
 - Develop EDA tools for 3DSoc compute/memory designs
 - Support tools for advanced 3DSoc designs

Metric	Goal
3DSoc Capability	> 50X 7nm 2D PaP
Hardware Accuracy	< 2% deviation from 3DSoc technology targets
Yield	> 30% for full 3DSoc designs
EDA Tools	Successful use of EDA flow for a > 500M gate/4GB memory design



www.darpa.mil



Drive pathfinding research efforts in new computing and communication technologies through enhanced DARPA-industry collaboration, funding, and guidance of university center research

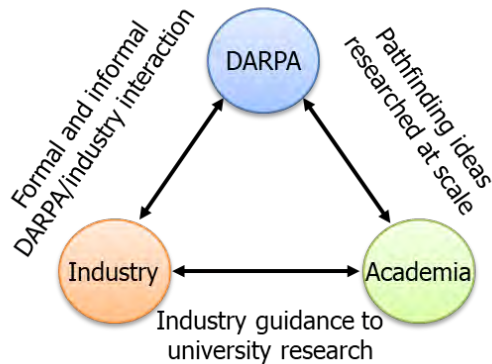


December 19th, 2018



Joint University Microelectronics Program (JUMP)

- Objectives
 - Drive long-term research in microelectronics with key players in industry and from academia
 - Develop long-range ideas that will drive formation of new DARPA programs
- Program overview
 - 6 centers focused on 6 major long-range microelectronic research themes
 - DARPA + 12 or more industrial sponsors
 - \$40M/year anticipated funding for 5 years (\$24M/year: industry and \$16M/year: DARPA)
 - US university faculty as PIs



JUMP Statistics

- 31 Universities
- 632 Students
- 129 Faculty Researchers
- 238 Liaison Personnel (Industry)

This Year

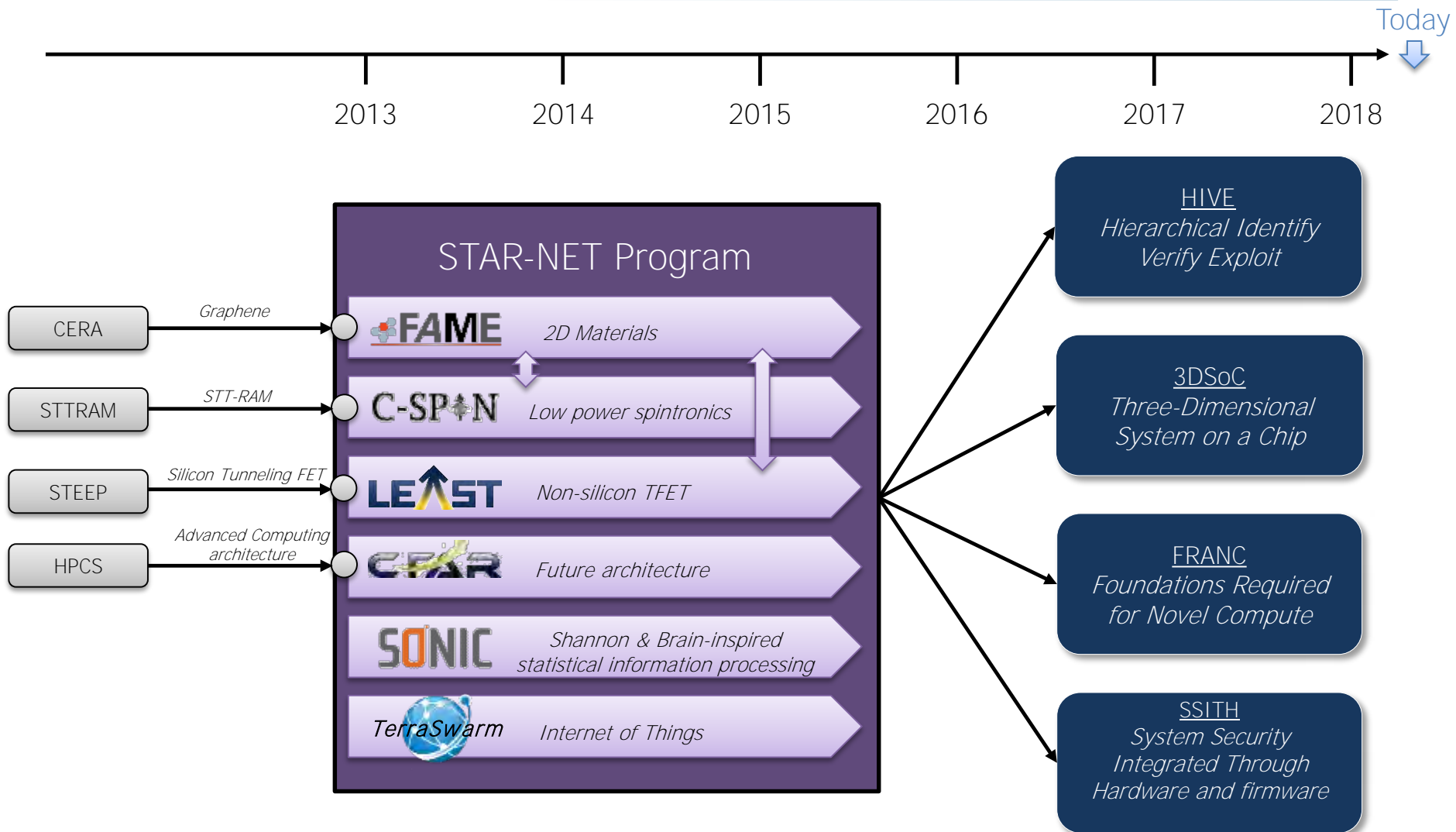
- 10 Contract starts
- 176 Task Starts
- 968 Research Publications
- 5 Patent Applications

JUMP Sponsors

DARPA	Intel	IBM
Raytheon	Micron	Analog Devices
Lockheed Martin	Samsung	TSMC
Northrop Grumman	ARM	EMD Performance Materials
	SK Hynix	



Path finding for future DARPA Programs



Systems/Applications



RF to THz
Sensors and
Communications



Intelligent
Memory and
Storage



Distributed
Computing and
Networking



Cognitive
Computing

Core Technologies



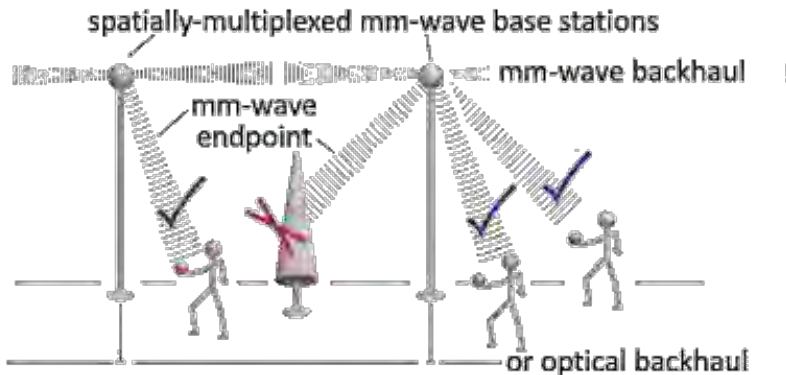
Advanced Architectures
and Algorithms



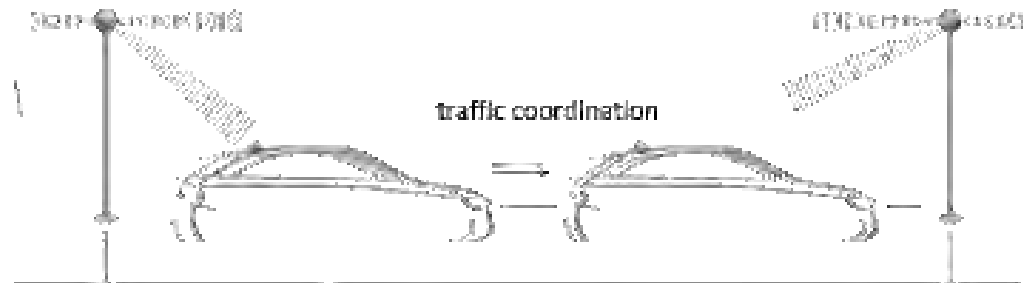
Advanced Devices,
Packaging, and Materials

Enabling 6G and Beyond

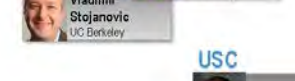
100+ GHz Communication



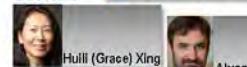
100+ GHz Sensing



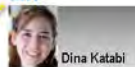
University of California



Cornell



MIT



NYU



Stanford



UT Dallas



USC



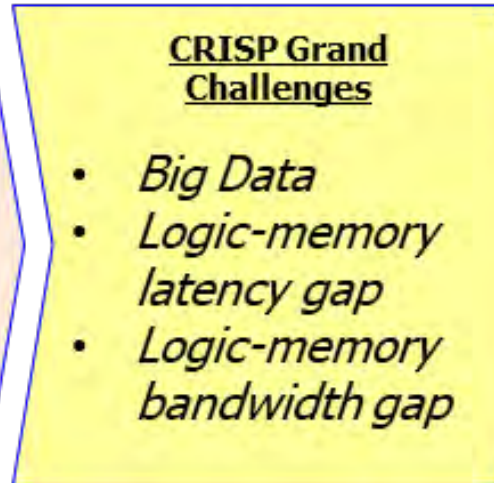
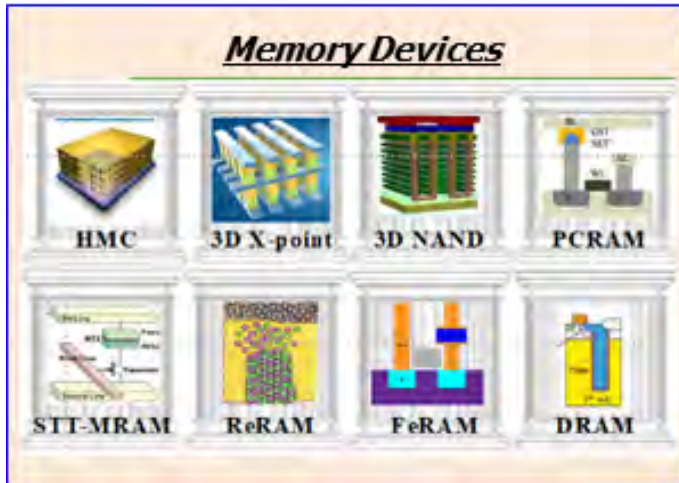
Theme 1: Systems and Algorithms for Converged THz Networks

Theme 2: mm-wave/THz ICS and Arrays for Communication and Imaging

Theme 3: Application-specific THz Transistors

Theme 4: Center-wide Demonstration Vehicles

Enabling Intelligent Memory



Univ. of Virginia



Samira Khan
Kevin Skadron
(Director)
Mircea Stan

Cornell



José Martínez
(Theme 1 Lead)
Zhiru Zhang

Penn State



Vijay Narayanan
Anand Sivasubramaniam

Illinois



Wen-mei Hwu
(new in July)

Stanford



Christos Kozyrakis
(new in July)

UCLA



Jason Cong

UC Santa Barbara



Dmitri Strukov



Yuan Xie
(Associate Director)

Univ. of Washington



Luis Ceze

UC San Diego



Rob Knight



Tajana Rosing



Steve Swanson
(Theme 2 Lead)



Jishen Zhao



Yuanyuan Zhou

Wisconsin



Jing Li



Jignesh Patel
(Theme 3 Lead)



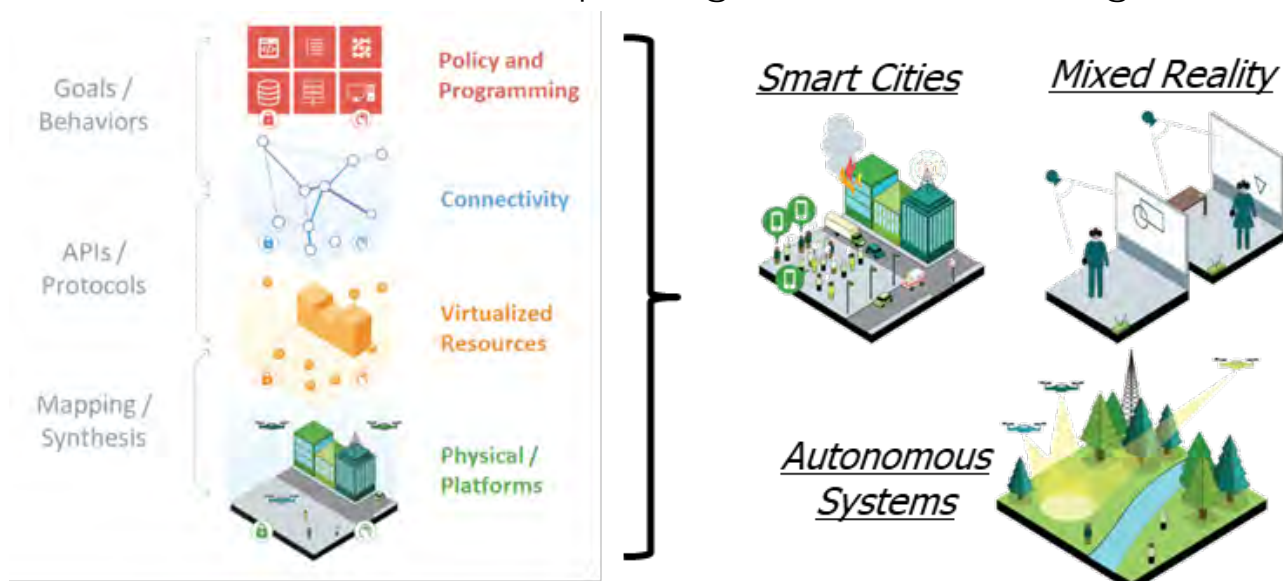
Kevin Elliceir

Theme 1: Hardware

Theme 2: Rethinking System-level Abstractions

Theme 3: Scaling Applications and Making the Programmer's Life Easy

Distributed Computing and Networking



UC Berkeley Prabal Dutta CO-DIRECTOR David Culler Jan Rabaey Claire Tomlin John Wawrzyniec	University of Washington Jeff Bilmes Ras Bodik 	Carnegie Mellon University Anthony Rowe DIRECTOR James Hoe Chris Harrison Vyas Sekar Brandon Lucia Bryan Parno
UC Los Angeles Mani Srivastava Paulo Tabuada	University of Southern California Ramesh Govindan Hao Li	UC San Diego Delan Stefan Rajesh Gupta

Theme 1: Physically-coupled Cognitive Perceptual System

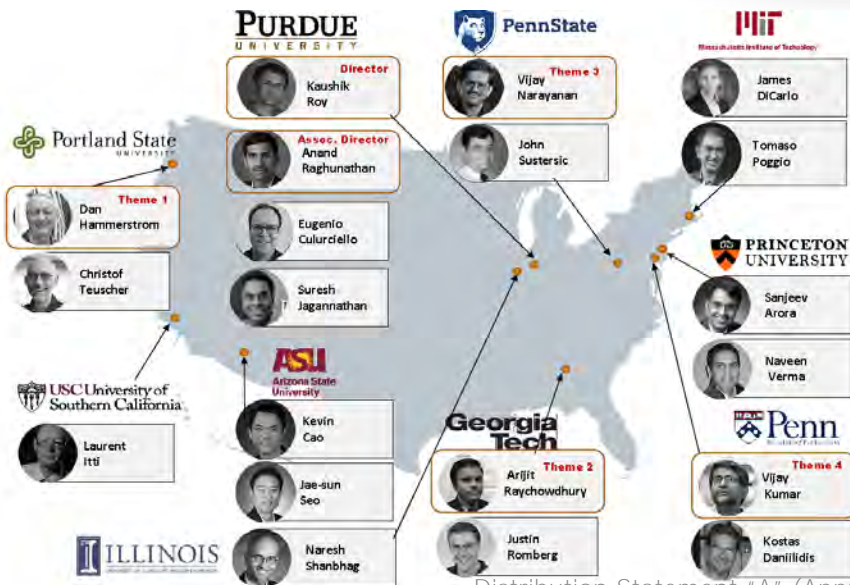
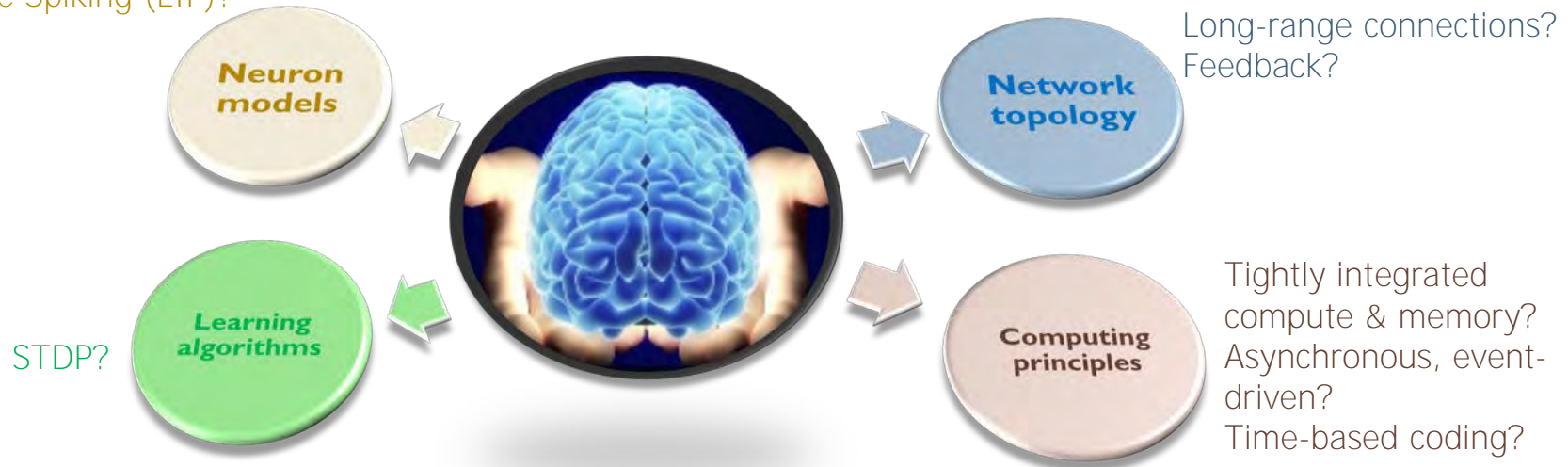
Theme 2: Platforms, Programming & Synthesis

Theme 3: Security, Robustness and Privacy

Theme 4: Interacting Services

Enabling Autonomous Intelligence

Discrete Spiking (LIF)?
HH?



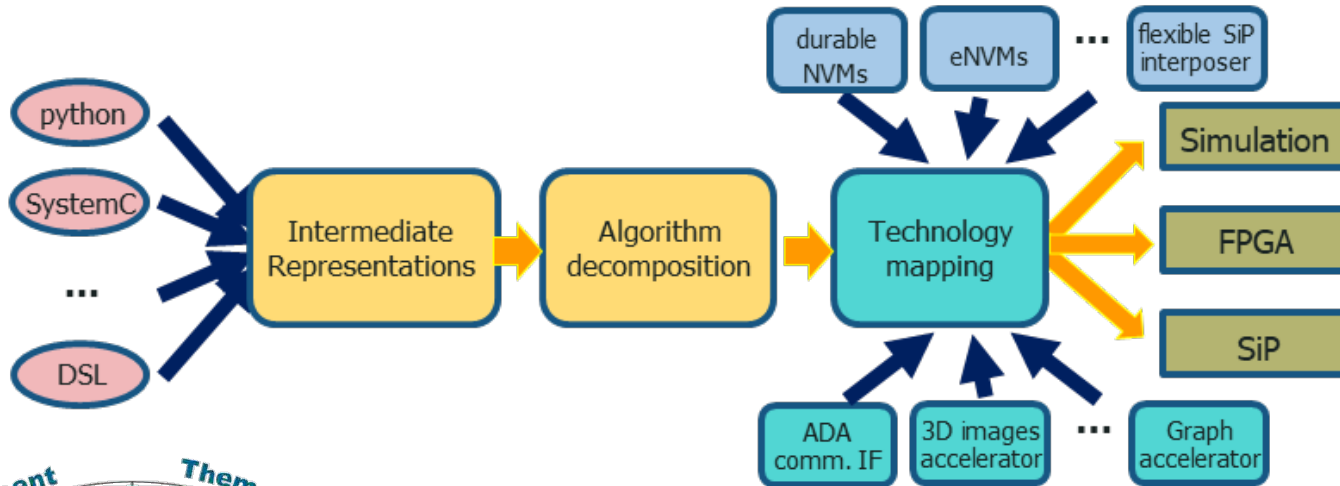
Theme 1: Neuro-inspired Algorithms and Theory

Theme 2: Neuromorphic Fabrics

Theme 3: Distributed Intelligence

Theme 4: Application Drivers

Enabling Low Cost and time-to-market Designs

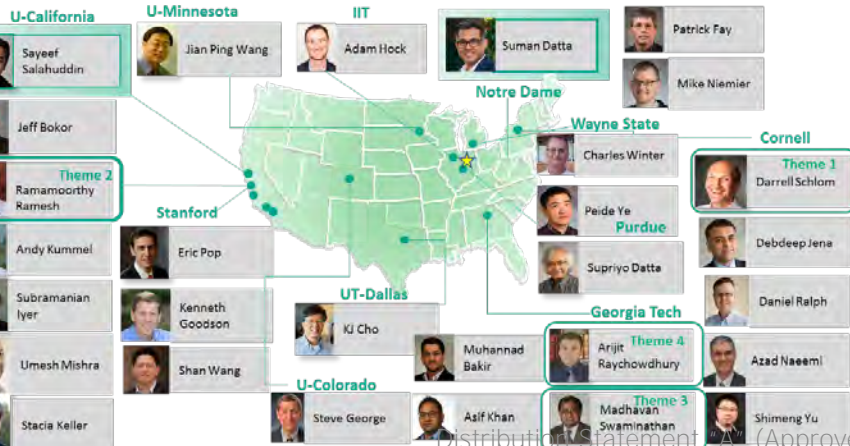
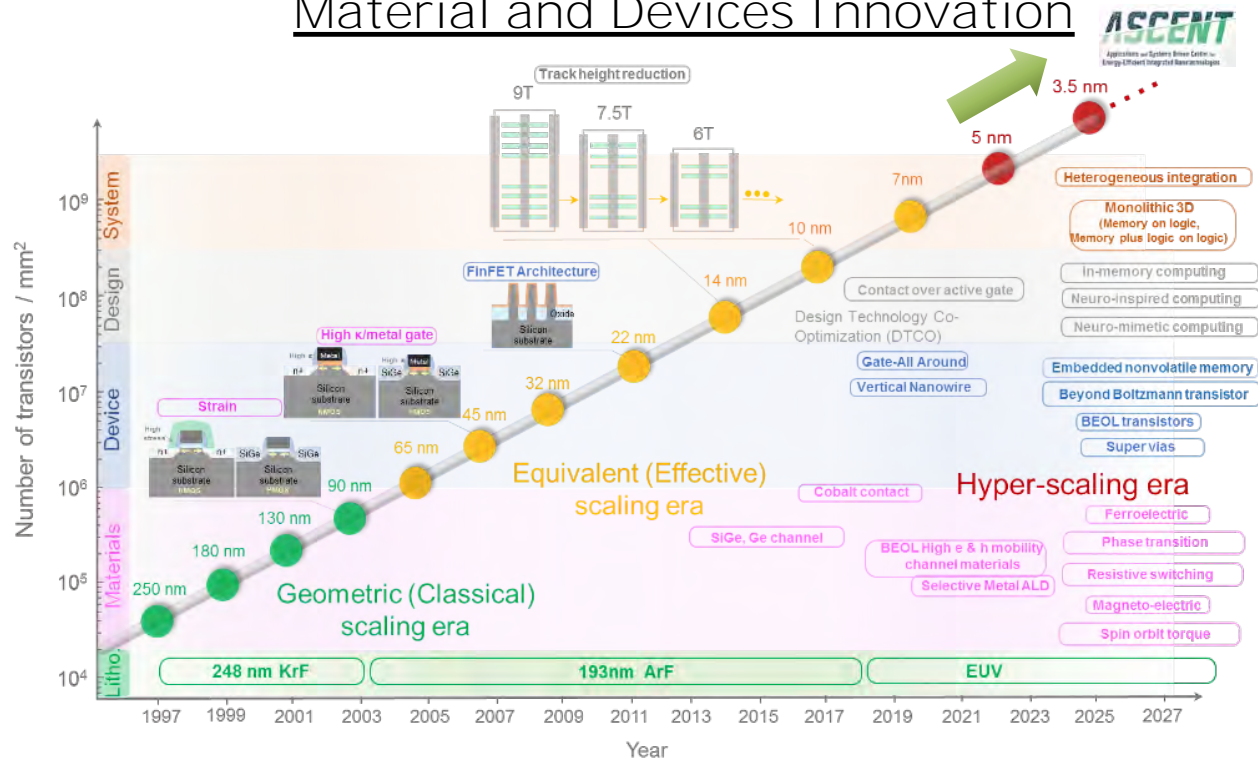


Theme 1: Agile System Development

Theme 2: Algorithm-Driven Architectures

Theme 3: Technology-Driven Systems

Material and Devices Innovation



Theme 1: Vertical CMOS

Theme 2: Beyond CMOS

Theme 3: Heterogeneous Integration Fabric

Theme 4: Merged Logic-Memory Fabric

*Let's collaborate to further develop JUMP
concepts into DoD applications!*



www.darpa.mil

linton.salmon@darpa.mil