Proposers’ Day Workshop
Monday, January 23, 2017

@srcJUMP, #JUMPpdf
Advanced Architecture and Algorithms

Ian Young
Senior Fellow,
Director Exploratory Integrated Circuits
Technology and Manufacturing Group
Intel Corporation
Center Motivation

ARCHITECTURE RESEARCH:
- Heterogeneous
- Distributed and Networking
- Cognitive and Non-traditional
- Intelligent Memory
- Communications Systems

ALGORITHMS and APPLICATIONS:
- Optimum algorithm / architecture
- Programming model
- Training structure
- Learning structure
- Computational complexity and scalability

Optimization
Co-Design
Architectures researched in;
- heterogeneous computing,
- non-traditional computing,
- distributed computing,
- intelligent memory,
- communications systems computing

Need to consider how the best algorithms can be implemented on these architectures
- to enable optimum performance, energy efficiency, scalability and cost.

Combine/co-design novel architectures with algorithmic implications imposed by computer science
• This theme must create theoretical foundations for programming and implementation of efficient algorithms on the new architecture paradigms within JUMP:
  ➢ Scalable, heterogeneous computing
  ➢ Cognitive computing
  ➢ Non-traditional computing
  ➢ Distributed Computing and Networking
  ➢ Communications Systems Algorithms
  ➢ Intelligent Memory

• Co-design considering the architecture and algorithmic implications.

• This theme will bridge the physical implementation of integrated circuit and architecture advances required to implement novel computation, communication, and storage applications, including those researched in the application-focused centers.
1) Architectures and algorithms using CMOS or beyond-CMOS technologies that are significantly lower in power and capable of utilizing
   • Up to $10^{12}$ devices per chip or
   • $10^{12}$ nodes per network
   • **100X lower energy**
   • Employing novel co-design to bridge the gap between architectures and algorithms for;
     • Optimization
     • Combinatorics
     • Computational geometry
     • Distributed systems
     • Learning theory
     • Online algorithms
     • Cryptography
   • Along with implementation, a strong component of theoretical computer science and modeling is expected.
   • The innovations that result from this foundational theme are expected to impact a broad variety of workloads and system applications.

2) Design and integration challenges of: systems composed of on-chip and off-chip accelerators, computation in and/or near data, and non-traditional computing.
New acceleration approach must demonstrate significantly higher system value, in terms of efficiency and/or performance, over known approaches such as CPU+GPU.

- Broadly-applicable architectures for compute in and near data (such as in-sensor or in-memory computing) are in scope.
- These advanced architectures should also provide mechanisms for workload-driven runtime composition of hardware resources, as well as auto-configuration and auto-tuning of system parameters.
- Reliability, resiliency, and energy-efficiency of application implemented on the new architectures must be addressed.

Centers focused on this area must enable scalable (heterogeneous) architectures using the relevant circuit primitives to realize significantly improved applications.
Theme Requires:

- Novel architectures should bridge Hardware to Algorithms Co-design.

- Benchmarking of the novel architectures.
  - Modeling and software innovations
  - To identify and remove barriers to hardware implementation or mass adoption.

Combine/co-design novel architectures with algorithmic implications imposed by computer science
Possible Research Tasks* (1)

1) Novel Architectures:
   - Overcome the connectivity bottleneck
   - Explore architectural mechanisms, microarchitectures, and alternative protocols to minimize communication and communication cost at all levels
     - In memory computing
     - Reconfigurable computing

2) Hardware-software co-design:
   - Design and integration of accelerators, from circuit level up to system & software level for heterogeneous systems, based on both existing and emergent devices (i.e. CMOS and Beyond CMOS)
   - Demonstrate dynamic control of operating points to
     - Manage the reliability of the computation
     - co-optimize Power and Performance
   - Develop non-conventional technologies for accelerators

* Possible research tasks of interest includes but are not limited to.
3) **System Integration:**
   - Develop software mechanisms to enable **transparent usage** of accelerators
   - Architectural and software innovations for enhanced **system flexibility and composability**, including for the integration of novel accelerators, allowing workload-driven composition of systems at runtime
   - Develop mechanisms for auto-configurations and auto-tuning of **heterogeneous systems**

4) **Software:**
   - **Software-defined infrastructure and resource virtualization**
   - Programming paradigms and languages for architectures using emerging technologies
   - Novel approaches to software engineering and developer operations
     - Including computer-supported programming / program synthesis

*Possible research tasks of interest includes but are not limited to.*
5) **Energy efficiency:**
   - Energy efficient circuits/architecture, algorithms and software: hardware/software/algorithm co-design methodologies and tools (includes underlying physics, physics of failure).
     - Required for energy efficient sensor systems, information extraction and autonomous systems
   - Real-time energy optimization. System-level design or circuit-based techniques that provide granular power gating. This may include feed-back or feed-forward with high- or low-resolution depending on the scenario or system.
   - Advanced **power management methods (e.g. non-volatile logic)**

6) **Scaling:**
   - Architectures that explicitly scale to $10^{12}$ logic devices per processor (not including memory)
   - **Fundamental understanding** of the new architectures that enable scaling up the number of devices
   - **Computational theory** of architectures that can use Beyond CMOS devices

---

*Possible research tasks of interest includes but are not limited to.*
5) **Memory related:**
   - Systems on chip with novel memory hierarchy
   - Nonvolatile computing (HW and SW)
   - **Computing on non-traditional fabrics** e.g. cross points

6) **Privacy and encryption:**
   - Accelerated *homomorphic encryption*
   - Authentication protocols and reliability supported by both HW and SW systems.
   - Computational theory for encryption and privacy

7) **Modeling of the new architectures: Verification and Validation**
   - Modeling – simulating the performance of the new heterogeneous architectures, **benchmarking** them against von Neumann machines
   - Modeling of performance: needs demonstration for a **broad enough class of workloads** and applications.
   - Modeling: validate modeling results against hardware performance.
   - Prove the **relevance** of the system hardware to **real-world applications**

*Possible research tasks of interest includes but are not limited to.*
While Centers are not limited to students from these disciplines, **students obtaining degrees** from the Advanced Architectures and Algorithms Center in the following areas of study are of particular interest to our sponsors in the years ahead:

- **MS / PhD**
- Applied Math
- **Comp Sci. & Comp. Eng.**
- **Electrical / Electronics Engineering or Electrical and Computer Engineering**
- Mathematics
- Systems Engineering