It’s CRNCH Time!

Summit 2024

Hyesoon Kim (CS), **Rich Vuduc** (CSE) — Co-Directors
Tushar Krishna (ECE) — Associate Director
Jeff Young (CS) — Rogues Gallery Director
Yingyan (Celine) Yin (CS) — Co-organizer for 2024
Fran Kendrick (CoC) — Administrative Support

Tom Conte (CS) CRNCH Founder
Vivek Sarkar (CS) — Co-Founder

Generous support by the College of Computing

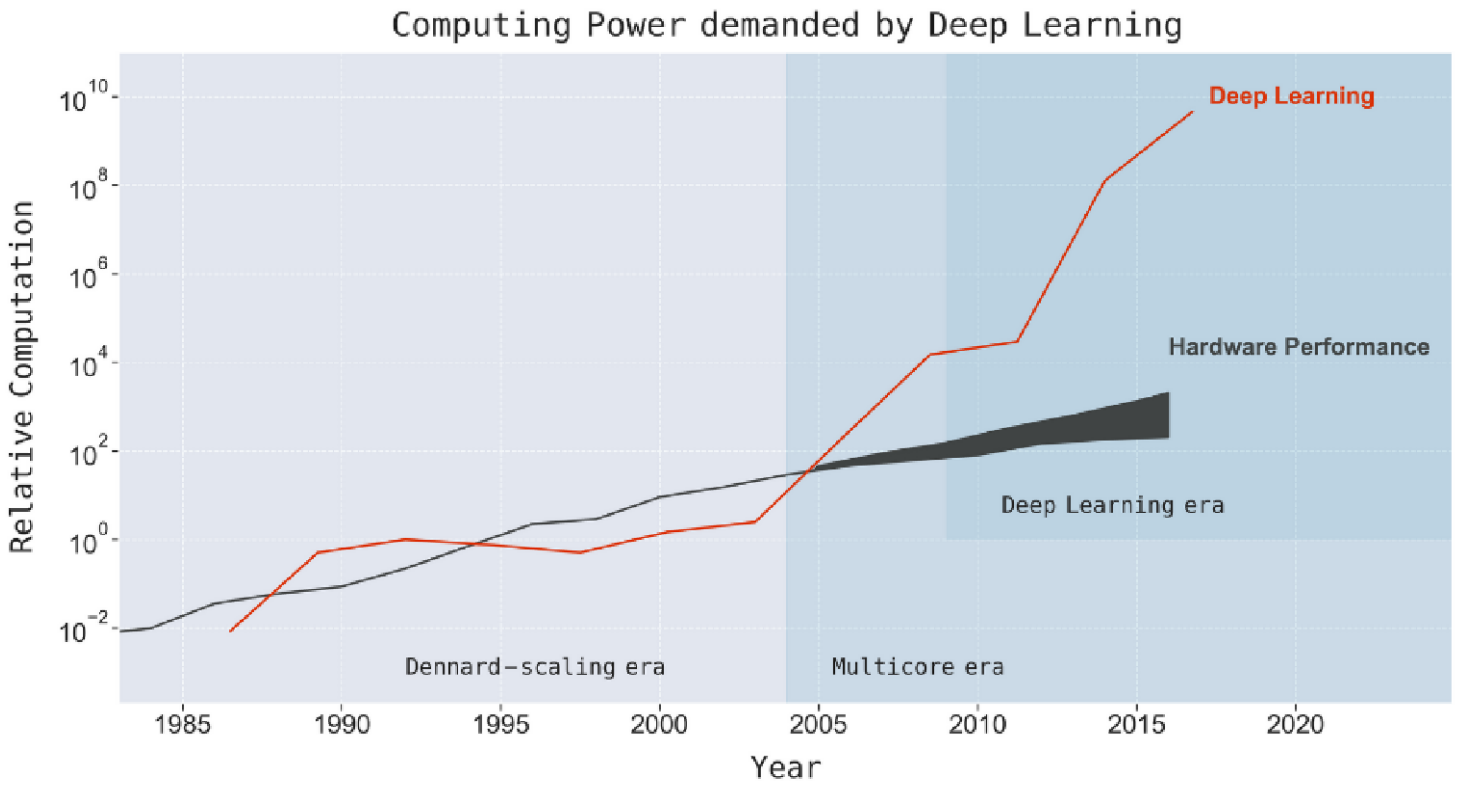
Georgia Tech College of Computing
Center for Research into
Novel Computing Hierarchies


You are CRNCH!
(Center for Novel Research on Computing Hierarchies)

Satiating the beast…

We live in an era of unprecedented, ”super-exponential” demand for computing power, which far outstrips supply even under a bullish view of Moore’s Law.

Thompson et al. (2020). *The computational limits of deep learning*. arXiv:2007.05558.



CRNCH: Computing Beyond the Conventional

- **Community**: Gathering researchers at GT across long-term post-Moore areas like neuromorphic, quantum, reversible, and approximate computing
- **Multidisciplinary**: Engaging people across sciences, engineering, computing, and GTRI; and connecting them with external partners like ORNL, Sandia,PNNL, Lucata, Northrup Grumman, …
- **Infrastructure & training**: Deploying unconventional computer hardware (e.g., Rogues Gallery)
- **Education**: Designing and promoting curricula aimed at helping the next generation “speak across the stack”

CRNCH Ph.D. Fellowship Winners — AY 2024

Amey Agrawal (ECE)

Llamero: Workload Adaptive Systems for Large Generative Language Model Inference

Advisor: Alexey Tumanov (CS)

Christopher Jawetz (ME)

Next-generation flow simulations using Quantum LBM

Advisor: Alexander Alexeev (ME)

Lakshmi Sathidevi (ECE)

An Automated Simulation and Evaluation Framework for Bonding Technologies in 3D FPGA

Advisors: Callie Hao & Sungkyu Lim (ECE)

Jinsun Yoo (ECE)

Eroica: A FaaS Framework for Geo-Distributed and Heterogeneous Edge Sites with Mobile Clients

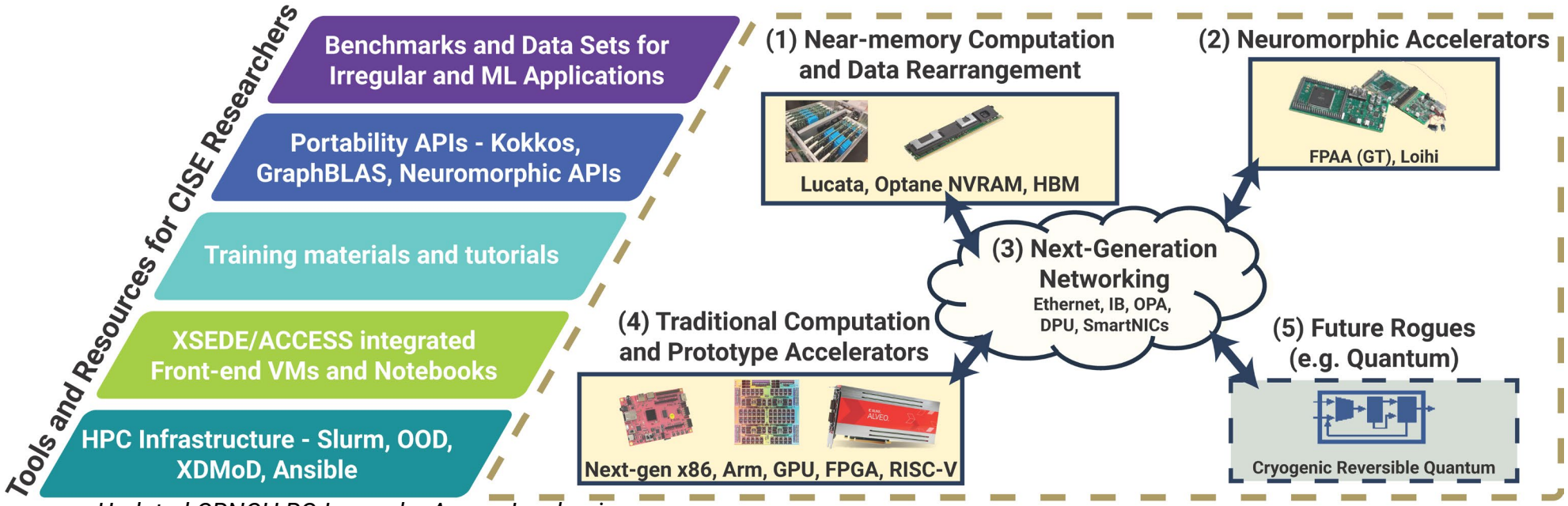
Advisor: Kishore Ramachandran (CS) & Tushar Krishna (ECE)

Rogues Gallery: Infrastructure for Post-Moore Computing Research (Lead PI: Jeff Young)

The Rogues Gallery is $1.3 million NSF project to create a testbed, including training materials, for the CISE research community.

At present, there are about 20 VMs, 30 servers, numerous boards, 150+ users (including 35 external), and over 120 students supported via a GT ”TechFee” (2021-2023).

Current hardware includes a rack-scale Lucata Pathfinder (16 nodes), neuromorphic accelerators; smart networking + 5G equipment; backend infrastructure; and novel chips and related benchmarking and testing.



Updated CRNCH RG Image by Aaron Jezghani

Forza Project ($15M iARPA Sponsored: Vivek Sarkar+)

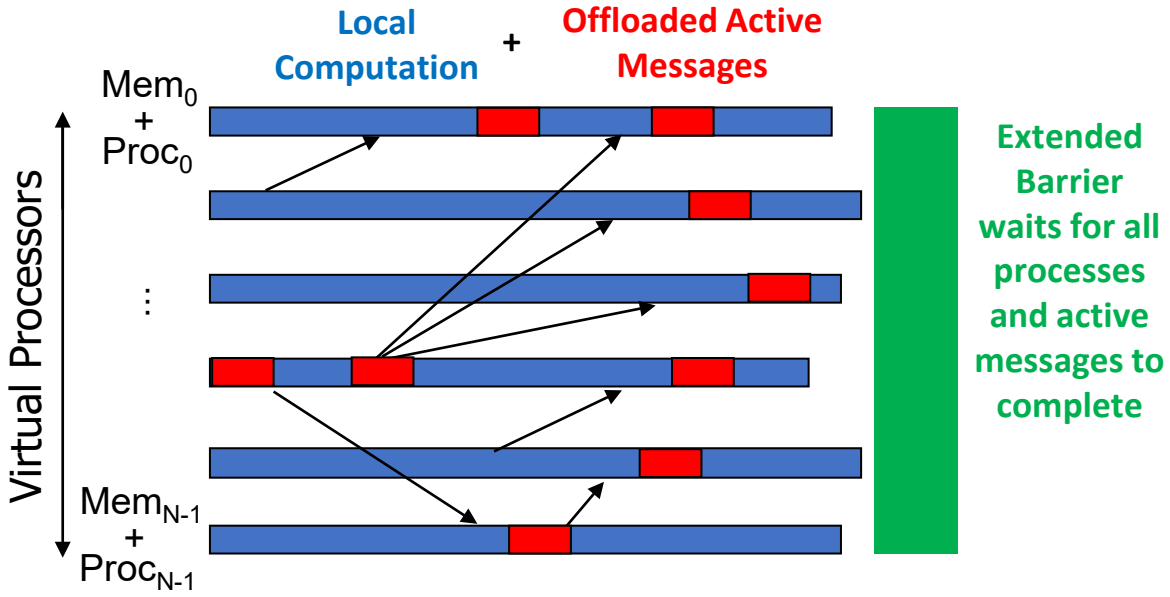
Software

Workflow-driven Co-design

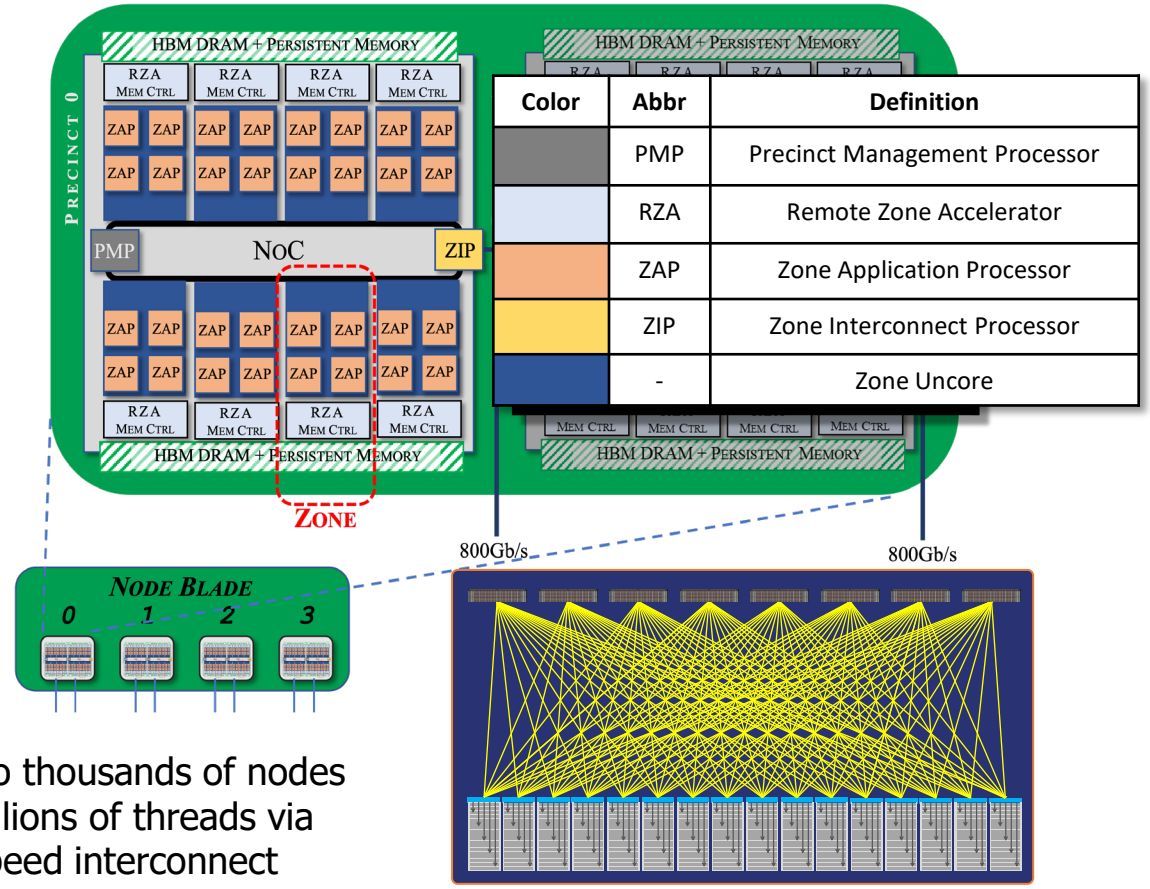
Hardware

Focus on approaches that *move compute to data via asynchronous active messages*

1. Zone Operations (ZOps)
2. Actor Messages
3. Functional threads (Migratable)



1 *precinct* ("0.25U node") contains 8 *zones*
1 zone contains 4 *ZAPs* ("cores")
1 ZAP *weave core* contains 512 *HARTs*



Color	Abbr	Definition
	PMP	Precinct Management Processor
	RZA	Remote Zone Accelerator
	ZAP	Zone Application Processor
	ZIP	Zone Interconnect Processor
	-	Zone Uncore

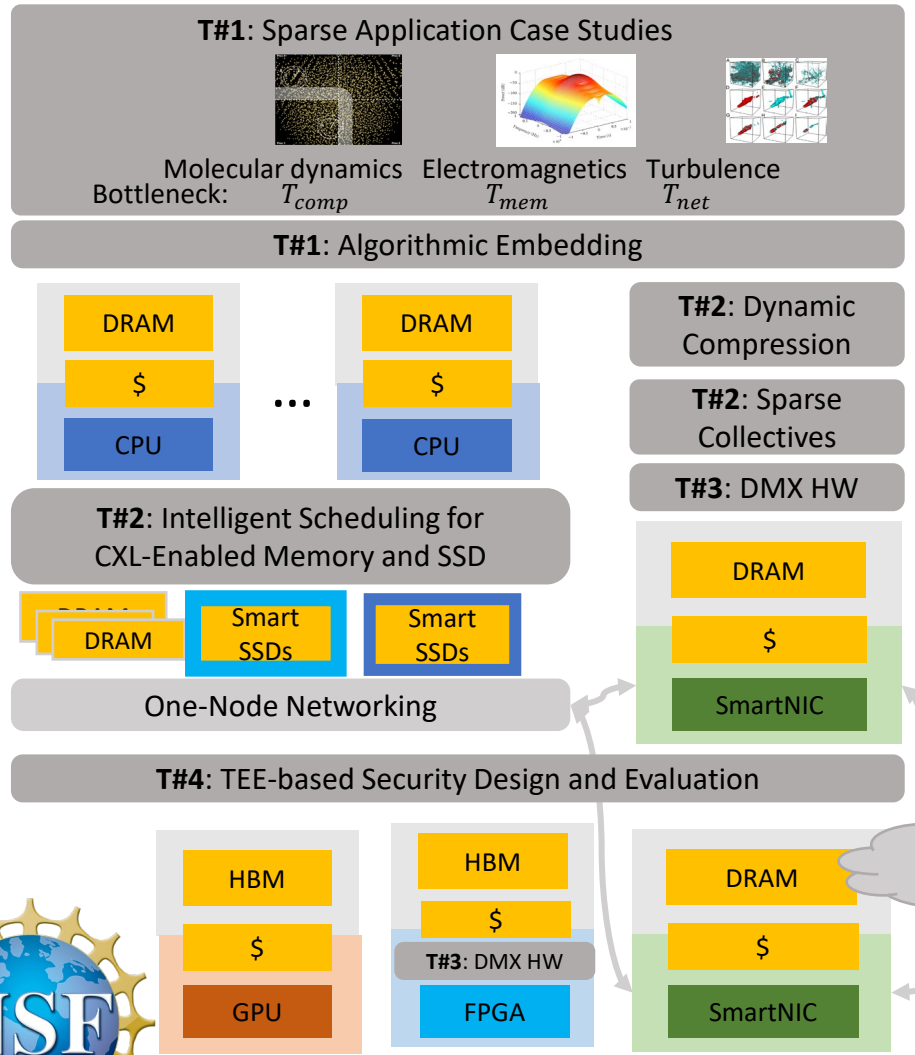
Scale to thousands of nodes and millions of threads via high-speed interconnect
(Full system = 1920 nodes in 16 compute racks + 4 switch racks)

FORZA

NSF 2316176, 2316177 PPoSS 2023: LARGE: ($5 million / 5 years)

Research into the Use and iNtegration of Data Movement Accelerators (RUN-DMX)

PIs: **Hyesoon Kim**, **Jeffrey Young**, **Richard Vuduc**, **Haesun Park**, **Bahar Asgari**



Goal: Shift away from accelerating compute to accelerating communication

Topics: Co-design of algorithms, software, and hardware co-design; utilizing smart NICs and smart SSDs; proposing new communication accelerators; developing and applying compression and near-data computation techniques

Task #1: Design Patterns for Algorithms and Applications (area: algorithm/theory) (1) Algorithm embedding, Sparse application case studies

Task #2: System abstractions and software (system/software): Intelligent scheduling for CXL Memory and SSDs, Dynamic compression

Task #3: Hardware co-design (architecture): DMX HW, including mathematical transformations, gate-level implementations of dependencies, fine-grain out-of-order execution, and on-the-fly sparsity-aware computations

Task #4: Security and Privacy (security) : TEE-based security design and evaluations

UNIVERSITY OF MARYLAND

Georgia Institute of Technology

Partner Highlights

CRNCH is grateful to its partners for enabling new efforts, so thanks, y'all!

RG expansion, research, tutorials, personnel transfers; CRNCH and Lucata were #46 on the Green Graph500 in 2021!

NORTHROP GRUMMAN

Novel computer architectures

Architecture, in-network computing, solvers, graphs, tensors, programming models

Just a (pseudo)random sample!

Summit 2024 — Themes

Co-design for ML … and beyond!

Brains, three ways — LLM, CS theory, analog

Novel infrastructures for computer systems-simulation

Scientific applications on new and future platforms

… and more!

Meet students at the afternoon poster session

crnch.gatech.edu