Arizona State University Center for Accelerated Real Time Analytics

ASU CARTA Capabilities

Ming Zhao, Associate Professor, ASU Site Director
School of Computing and Augmented Intelligence
Web: https://carta.asu.edu, Email: carta@asu.edu

Arizona State University

Interdisciplinary Faculty Team

Computer Science & Engineering

Dr. Ming Zhao

Dr. Fengbo Ren

Dr. Chris Bryan

Dr. Jia Zou

Industrial Engineering

Dr. Feng Ju

Dr. Hao Yan

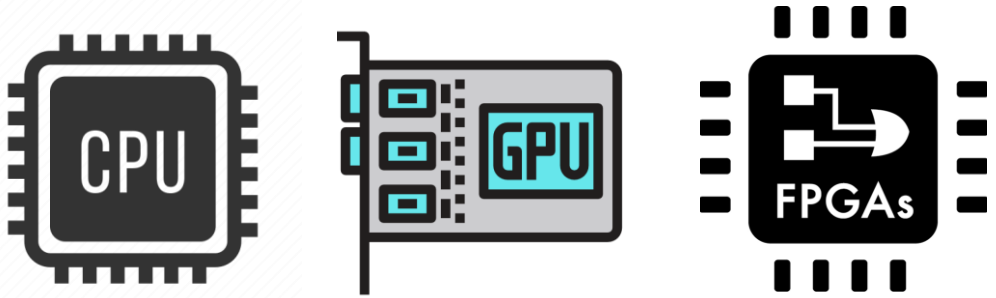
Software Engineering

Dr. Srividya Bansal

Acceleration & Heterogeneous Computing

- Heterogeneous infrastructure
  - Hardware accelerators (400 CPUs, 32 GPUs, 32 FPGAs)
  - Memory technologies (1TB DRAMs, 8TB persistent memories)
  - High-speed network (100Gb/s InfiniBand)

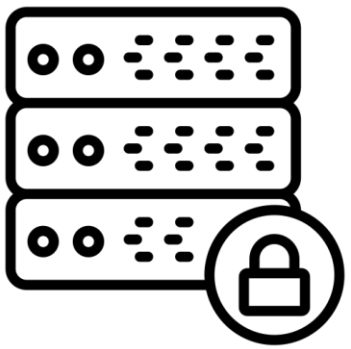
- Heterogenous computing
  - FPGA designs for accelerating deep neural networks & multidisciplinary applications (medicine, transportation)
  - Cross-platform computing using OpenCL

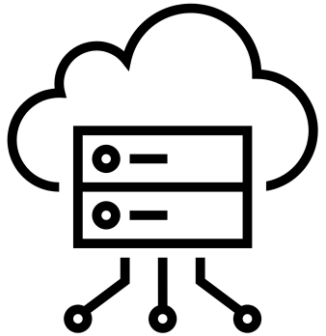


Real-time Edge Computing

- IoT and edge computing infrastructure
  - 150 Raspberry Pis, NVIDIA edge GPUs, Coral edge TPUs, Intel NUCs, Amazon DeepRacers, SparkFun JetBots
  - 50 Intel cloudlet nodes
  - Robotics lab
  - Manufacturing lab

- Real-time IoT-data-driven learning & analytics
  - Distributed machine learning
  - Machine learning model compression





Research Themes

- High-performance Computing
- Visual Analytics
- Deep Learning
- Real-time Data Analytics
- Data Fusion & Integration
- Smart Manufacturing

Heterogeneous High-performance Computing

Dr. Zhao & Dr. Ren

Research Objective

- Enable portability
- Enable scalability
- For both existing & future accelerators

Key Innovation

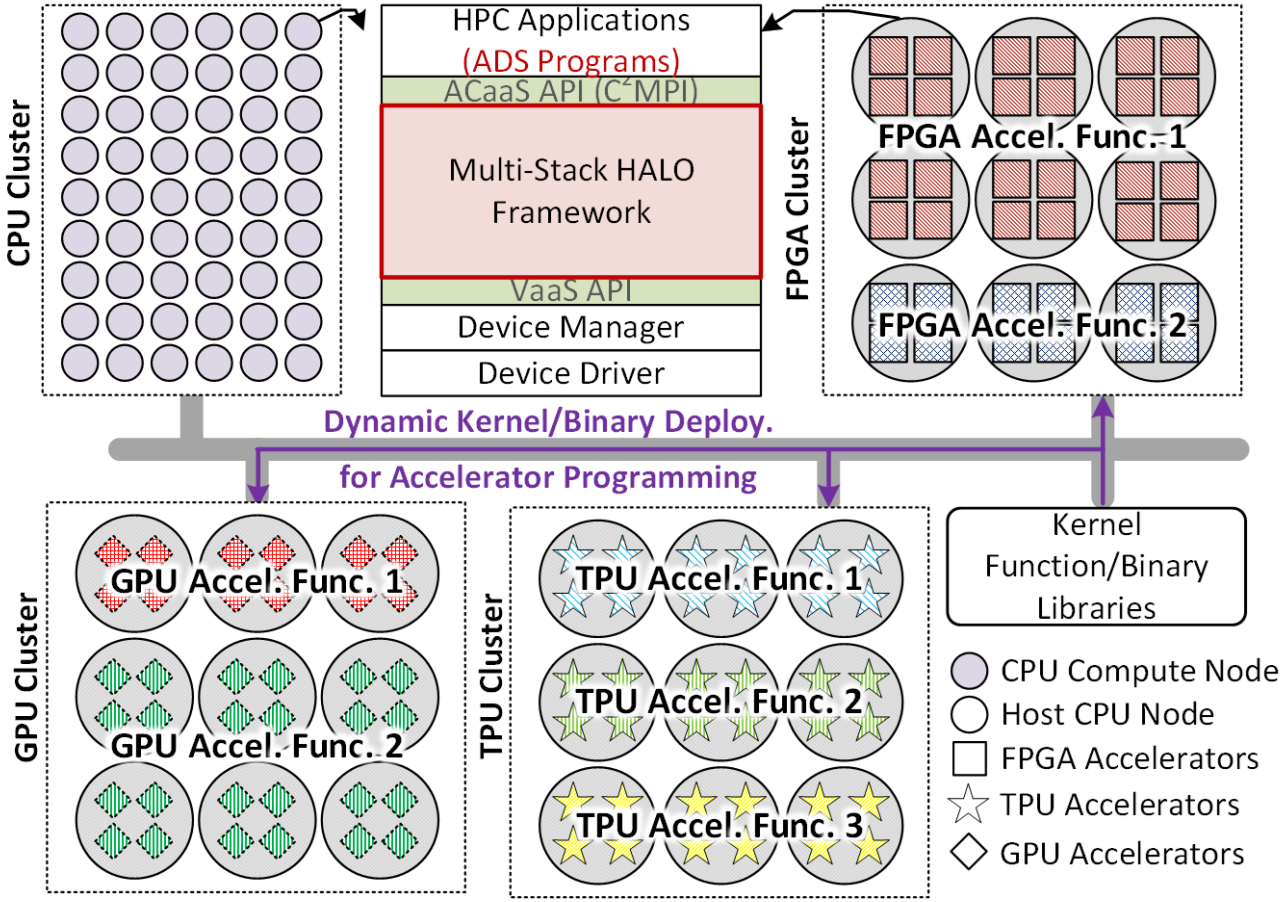
- Hardware-agnostic accelerator orchestration
- Compute-centric MPI

Broader Impact

- Effortless adoption of accelerators by domain experts

High Potential Sponsor

cadence



Distributed Deep Learning

Dr. Ming Zhao

Research Objective

- Accurate, efficient, and responsive learning for IoT data driven applications

Key Innovation

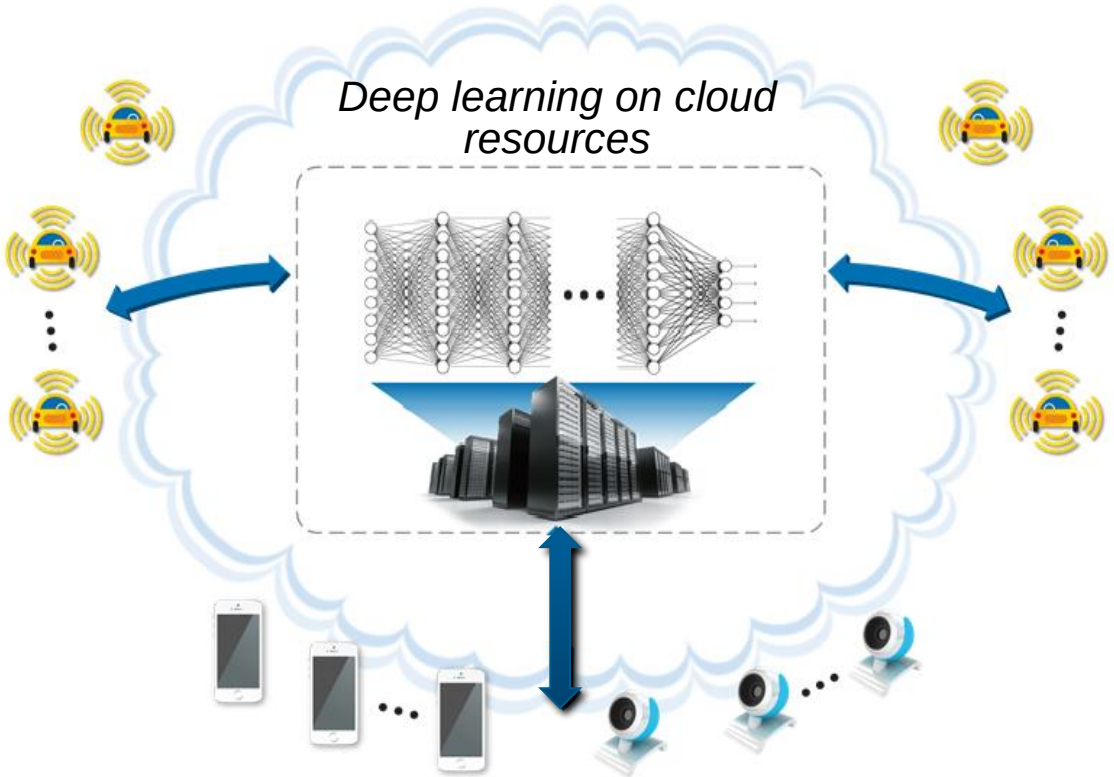
- Policy-based automatic model customization
- Cross-edge-cloud collaborative learning

Broader Impact

- Support for diverse data-driven, learning-based applications and heterogeneous deployment platforms

High Potential Sponsor

aws intel



Responsive/customized/private learning on the edge

Visual Analytics for Complex Data

Dr. Chris Bryan

Research Objective

- Designing novel visualizations and interfaces to supporting analysis of complex data

Key Innovation

- “Human-in-the-loop” processes combine human reasoning with data-driven models

Broader Impact

- Tools and techniques are widely applicable across a variety of data domains and problems

High Potential Sponsor

Adobe SRP Uber

Visualizing how anomaly events egocentrically propagate from a source location through the electric power grid network

Real-time Deep Learning Inferences from Relational Databases

Dr. Jia Zou

Research Objective

- To provide fast, low-cost, and scalable deep learning inference from relational databases

Key Innovation

- Automatic Tensor-Relational Translation
- Co-optimization of data serving and DNN model serving in one system

Broader Impact

- Reduce inference latency and costs for a broad class of AI + database applications

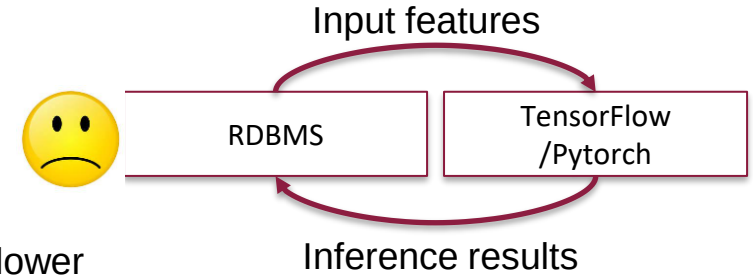
Current Sponsor

We are in contact with more potential sponsors

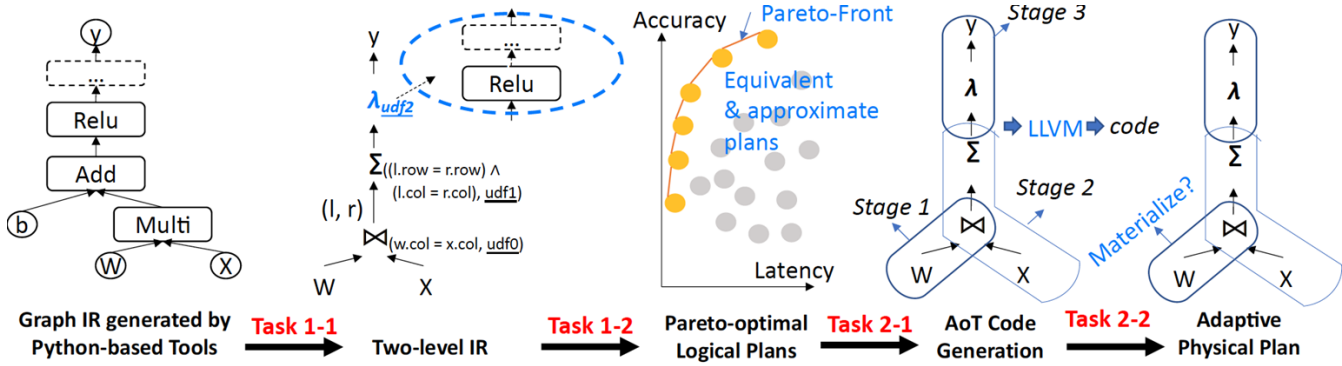
- ***Use cases:*** Smart transaction processing on relational data

IF (**fraud-model1.infer**(features(*transaction-id*)) IS FALSE):
Update ACCOUNT SET bal=*bal-amount* WHERE ID=*id1*
Update ACCOUNT SET bal = bal + amount WHERE ID=*id2*

- ***Existing approach:*** High latency, low resource utilization, high cost



- ***Proposed approach:*** **10x** lower latency and **4x** better resource utilization, significantly lower cost



Thank you!

http://carta.asu.edu
Email: carta@asu.edu

Arizona State University