ARA Distributed Inference Experiments

Flying HTCondor Over a Field of Wireless Dreams

Marty Kandes[1,3]

Sarath Babu[2,3] Zhibo Meng[2,3] Hongwei Zhang[2,3]

[1]San Diego Supercomputer Center

[2]Iowa State University

[3]https://icicle.ai

Throughput Computing 2025
University of Wisconsin - Madison
Howard Auditorium, Fluno Center
Thursday, June 5th, 2025
09:25 - 09:45 CT

ARA Distributed Inference Experiments

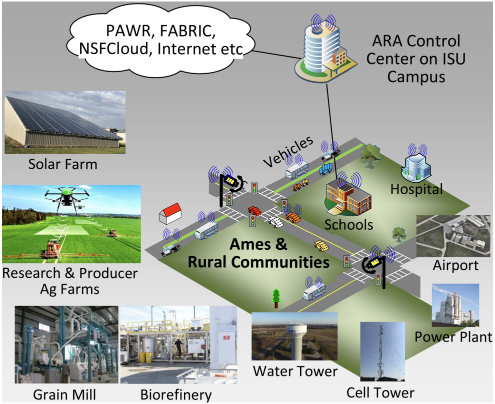
Flying HTCondor Over a Field of Wireless Dreams

ARA Wireless Living Lab

- One of the U.S. National Science Foundation (NSF) Platforms for Advanced Wireless Research (PAWR).
- Technology focus on long-distance, high-capacity wireless backhaul and radio access network (RAN) platforms for rural broadband.
- Deployed and managed by the Center for Wireless, Communitites, and Innovation (WiCI) at Iowa State University (ISU).
- ARA is an acronym for **A**griculture and ru**RA**l communities.

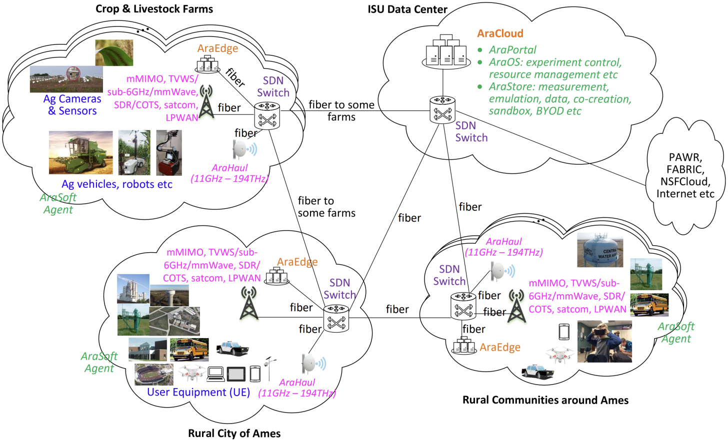
https://arawireless.org

ARA Vision



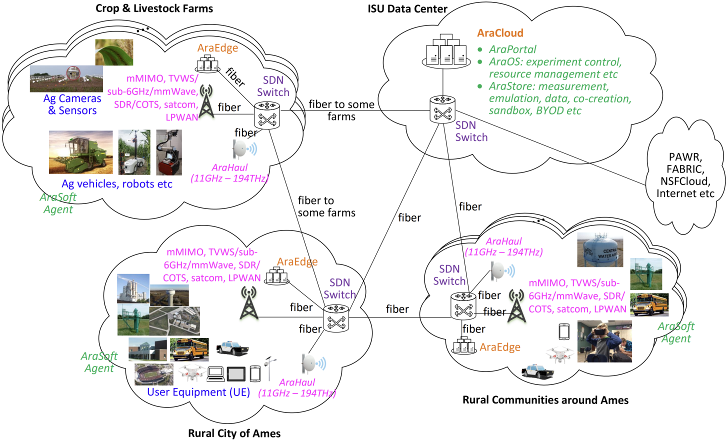
- Experiment with advanced wireless technologies at-scale.
- Foster innovation in wireless solutions and applications.
- Address future challenges and opportunities in rural broadband. e.g., precision agriculture; community services

ARA System Overview



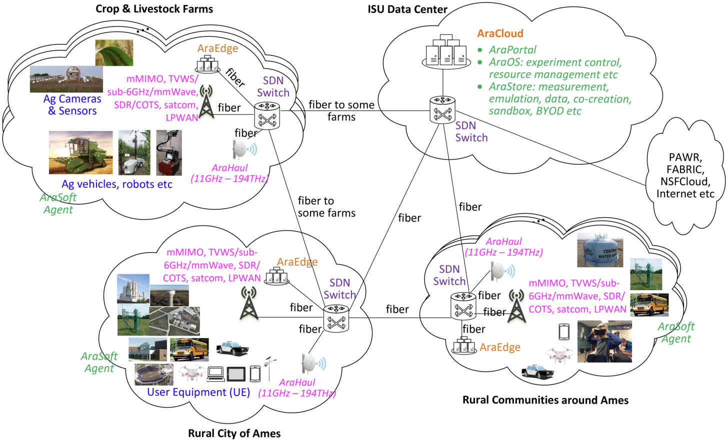
AraHaul is a multi-modal, multi-hop, wireless backhaul (x-haul) network for long-distance, high-capacity, point-to-point communication between nearby rural communities, farms, and the core ARA network at ISU.

ARA System Overview



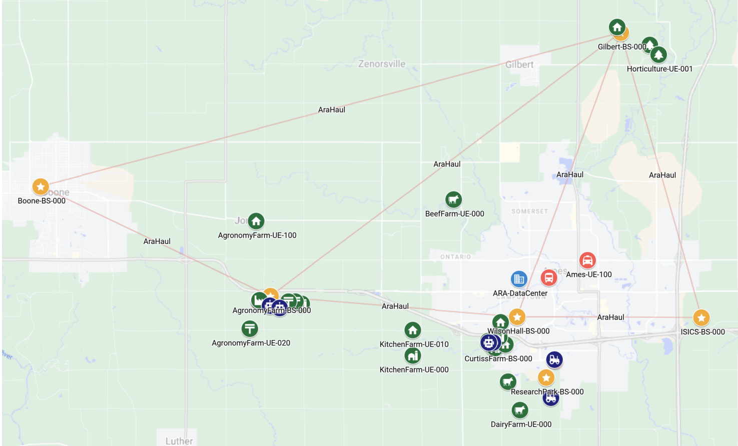
AraRAN is a heterogeneous radio access network (RAN) for high-capacity, low-latency wireless communication between user equipment (UE) in the field and base stations (BS) at *AraHaul* sites.

ARA System Overview



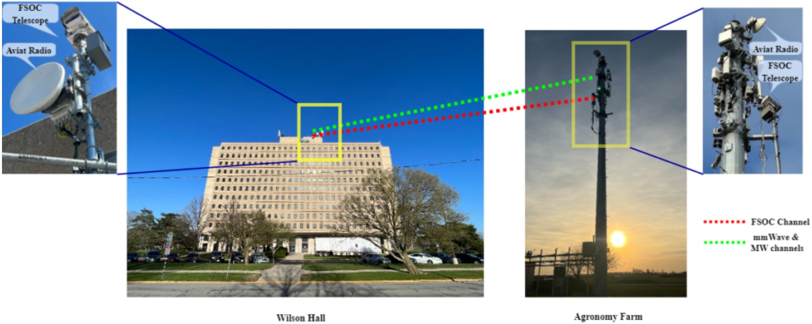
AraCloud/AraSoft is an OpenStack-based platform that controls and manages access, configuration, scheduling, and provisioning of the resources deployed across ARA to support experimentation by researchers.

Current ARA Deployment



ARA's Phase 2 deployment currently spans approximately 500 km^2 near ISU, including campus, the City of Ames, local research and producer farms, and neighboring rural communities in central Iowa.

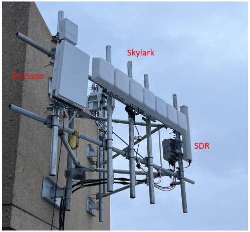
AraHaul Platforms



	Band	Frequency	Wavelength	Range	Capacity
AraOptical (FSO)	Near-IR	191.7 - 194.8 THz	1539 - 1564 nm	15 km	16 x 10 Gbps
Aviat WTM 4811	mmWave	81 - 86 GHz 71 - 76 GHz	3.49 - 3.70 mm 3.95 - 4.22 mm	15 km	1 x 10 Gbps
	μWave	10.6 - 11.5 GHz	2.61 - 2.83 cm	20 km	2 x 1.5 Gbps
Hughes HL1120W Eutelsat OneWeb	μWave	14.0 - 14.5 GHz 10.7 - 12.7 GHz	2.06 - 2.14 cm 2.36 - 2.80 cm	1200 km LEO	32 Mbps (up) 195 Mbps (down)

AraHaul: Multi-Modal Wireless X-Haul Living Lab
for Long-Distance, High-Capacity Communications

AraRAN Platforms: AraMIMO, AraSDR



Base Stations	Band	Frequency	Wavelength	Range	Capacity
Ericsson AIR 5322 B261	mmWave	27.5 - 27.9 GHz	10.8 - 10.9 mm	500 m	3 Gbps
Ericsson AIR 6419 B77G	μWave	3.45 - 3.55 GHz	8.5 - 8.9 cm	8.5 km	650 Mbps
Ettus USRP N320	μWave	3.4 - 3.6 GHz	8.3 - 8.8 cm	1.2 km	100+ Mbps
Skylark Faros v2	TVWS	539 - 593 MHz	0.51 - 0.56 m	10 km	100+ Mbps

User Equipment	Frequency	Wavelength	Capacity
Quectel RG530F	600 MHz - 28 GHz	10 mm - 0.5 m	3.4/8.8 Gbps (up/down)
Ettus USRP B210	70 MHz - 6 GHz	50 mm - 4.3 m	1.5 Gbps
Skylark Faros v2 CPE	100 MHz - 6 GHz	50 mm - 3.0 m	40/250 Mbps (up/down)

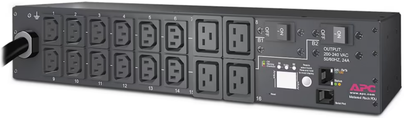
AraMIMO: Programmable TVWS mMIMO Living Lab for Rural Wireless

AraSDR: End-to-End, Fully-Programmable Living Lab for 5G and Beyond

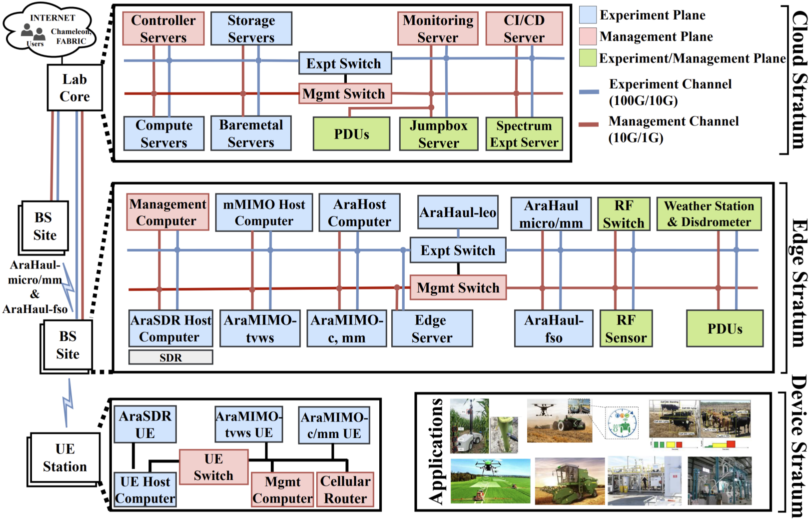
Other ARA Infrastructure

Agronomy Farm

Wilson Hall

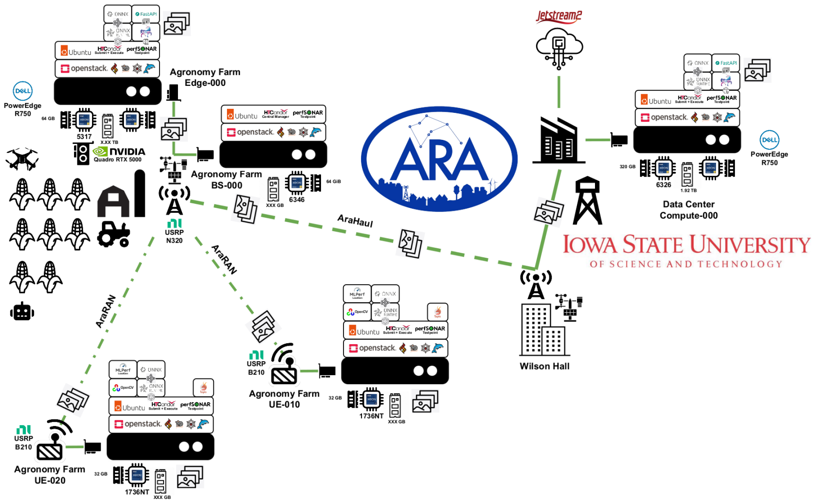


ARA System Architecture



Design and Implementation of ARA Wireless Living Lab for Rural Broadband and Applications

ARA Distributed Inference Experiments (ADIx)



To demonstrate, characterize, and evaluate the use of distributed inference for computer vision tasks in rural and remote regions ...

ADIx Design: MLPerf Inference Benchmarks

MLPerf

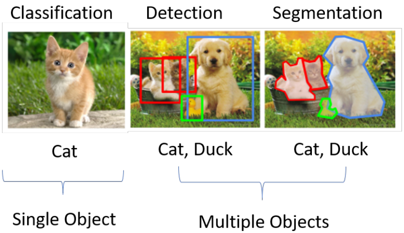
ML Commons

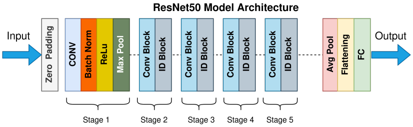
https://mlcommons.org

MLPerf aims to be a representative benchmark suite for AI/ML that evaluates system performance to meet five high-level goals:

- Enable fair comparison of competing systems while still encouraging innovation.
- Accelerate progress through fair and useful measurement.
- Enforce reproducibility to ensure reliable results.
- Serve both the commercial and research communities.
- Keep benchmarking effort affordable so all can participate.

ADIx Design: MLPerf Inference Benchmarks



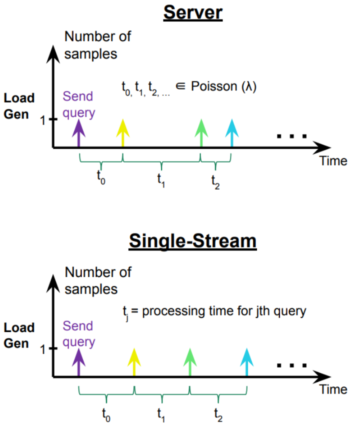


IMAGENET

All benchmarks measure how fast systems can process inputs and produce results using a trained model. Each benchmark specifies:

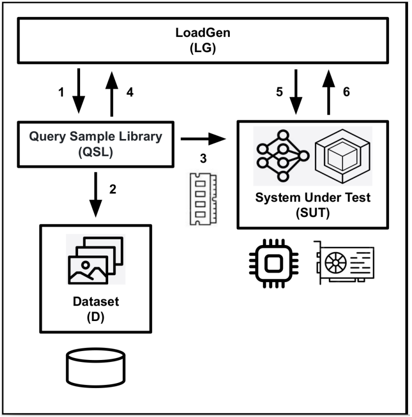
- **Task**: Image Classification, Object Detection
- **Model**: Resnet50-v1.5, Retinanet
- **Dataset**: ImageNet (224×224), OpenImages (800×800)
- **Quality**: 99% of FP32 (76.46%), 99% of FP32 (0.3755 mAP)

ADIx Design: MLPerf Inference Benchmarks



MLPerf also specifies the required inference workload **scenarios** and how they must be implemented and measured.

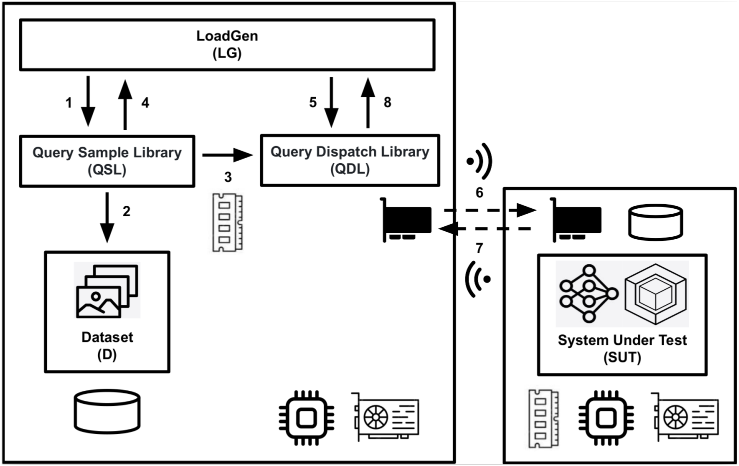
ADIx Design: MLPerf Inference Benchmarks (Closed)



- **LoadGen (LG)**: Generates traffic for scenarios (with QSL index)
- **Query Sample Library (QSL)**: Loads (dataset) samples into memory
- **System Under Test (SUT)**: Hardware and software to be measured

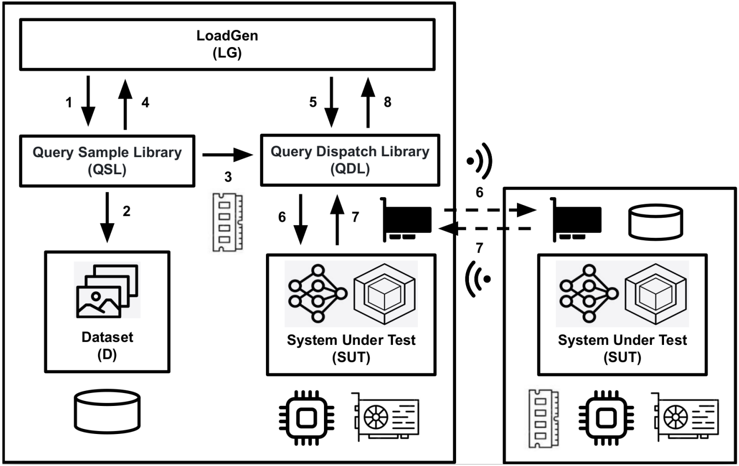
https://docs.mlcommons.org/mlcflow

ADIx Design: MLPerf Inference Benchmarks (Network)



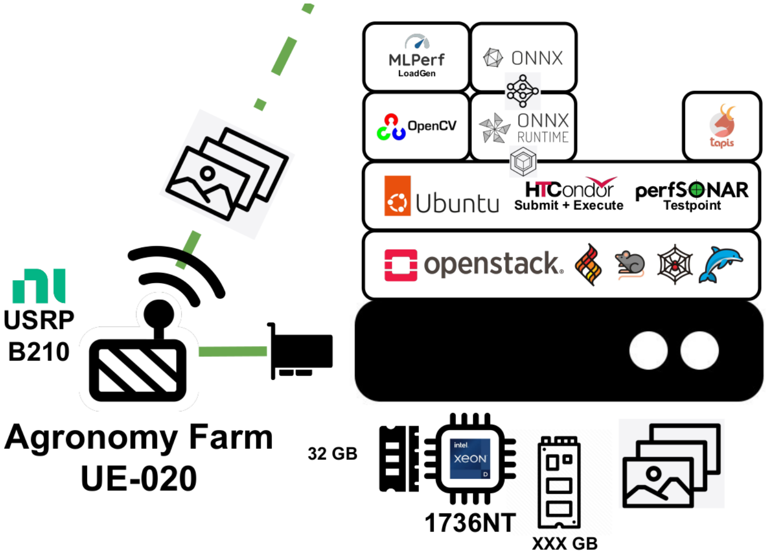
- **Query Dispatch Library (QSL)**: SUT proxy that runs on the LG system

ADIx Design: MLPerf Inference Benchmarks (Open)

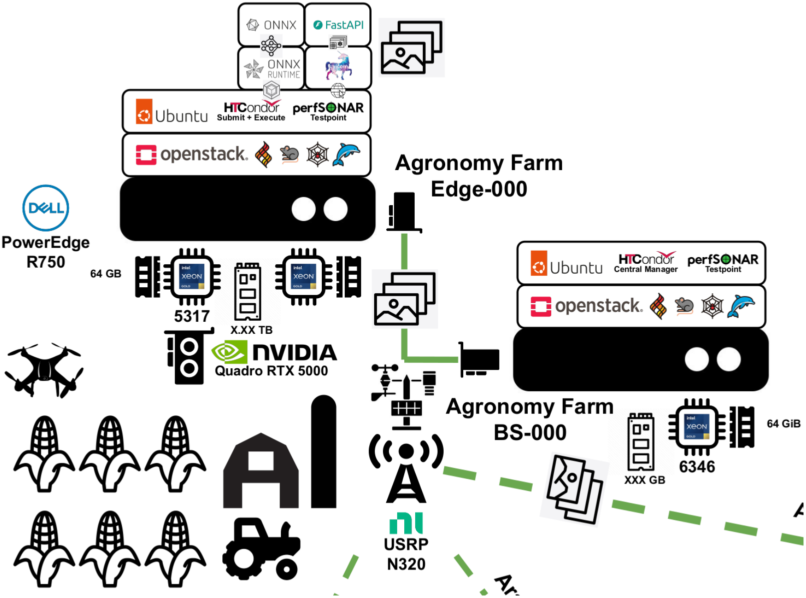


If we break the rules, there is more compute to be had.
What then might be possible?

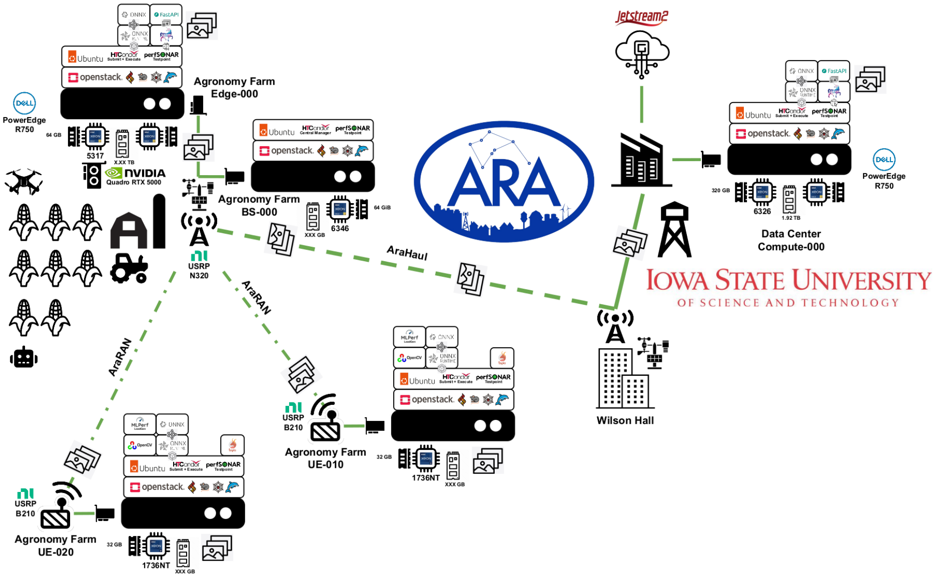
ADIx Implementation



ADIx Implementation



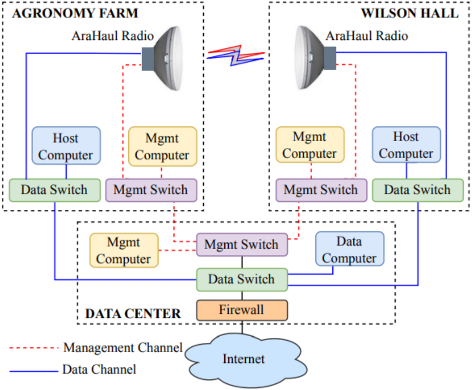
ADIx Implementation



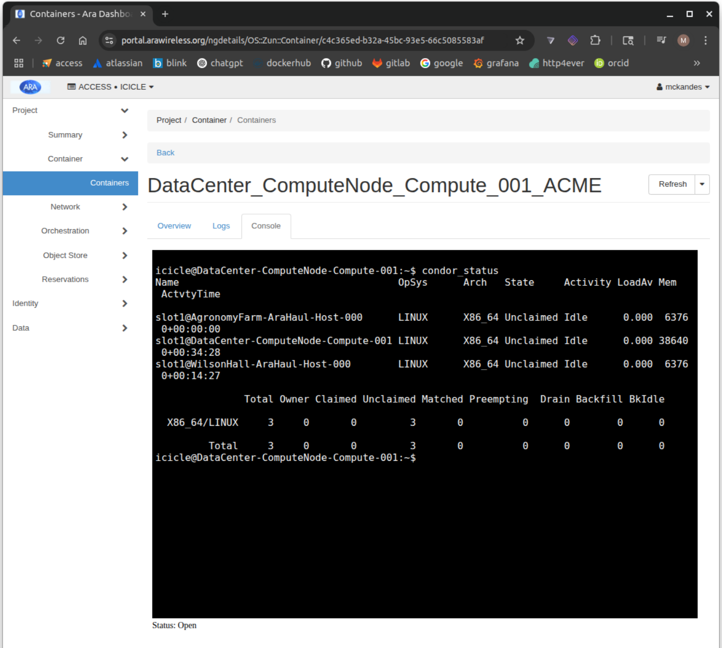
ADIx Implementation: Key HTCondor Features

- **CONDOR_HOST**: By default, ARA dynamcially allocates management IPs to resources each time you spin up a new experiment configuration; one IP need to know is helpful
- **Python bindings** imported as part of QDL to implement: (1) a direct high-thoughput batch (HTB) mode that distributes QSL samples to be processed by ORT-based SUTs; (2) warm start of FastAPI-based (ORT-backed) SUTs
- **Multi-homed support**: Most ARA systems will have multiple network interfaces in any given ADIx configuration
- **Condor Connection Broker**: Leverage compute resources beyond ARA in the future (e.g., Voyager's Habana inference nodes)

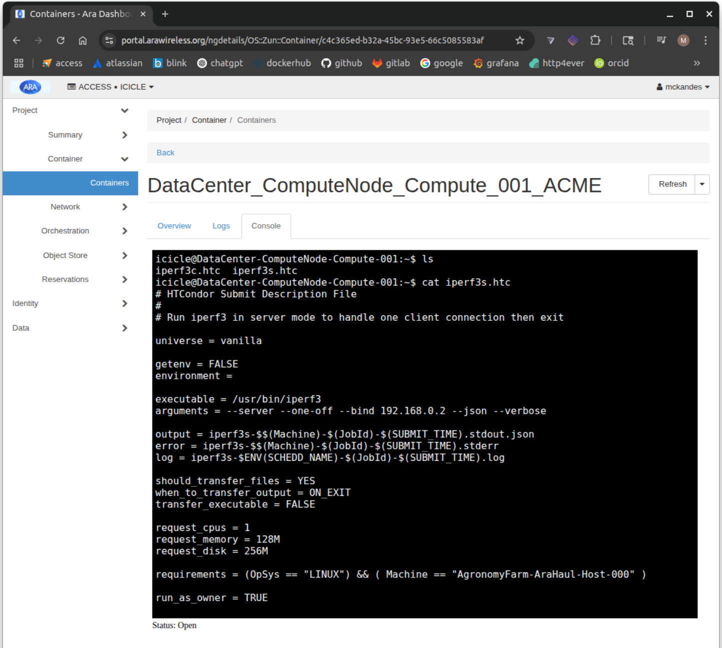
AraHaul Demonstration (with HTCondor)



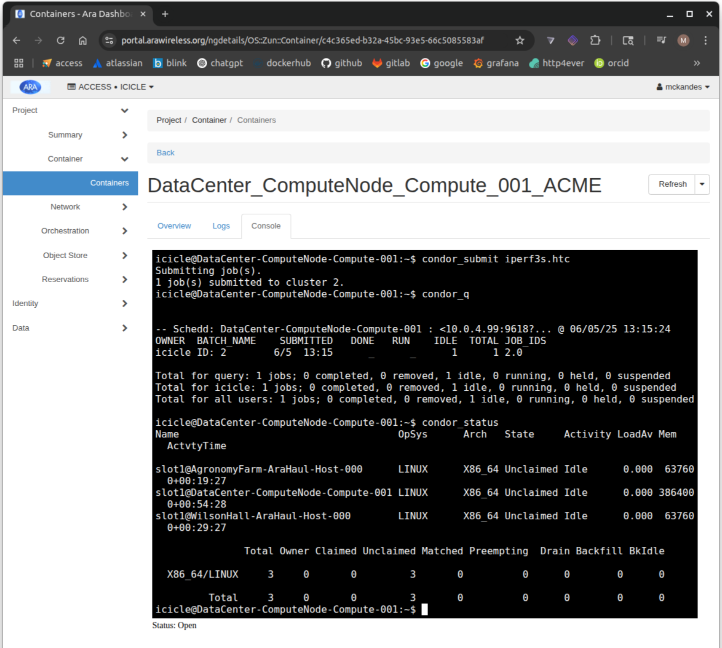
AraHaul Demonstration (with HTCondor)



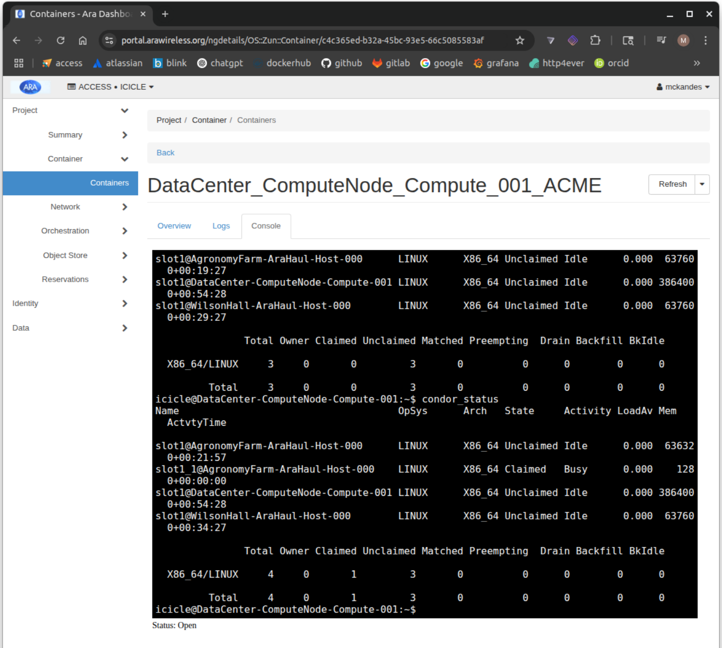
AraHaul Demonstration (with HTCondor)



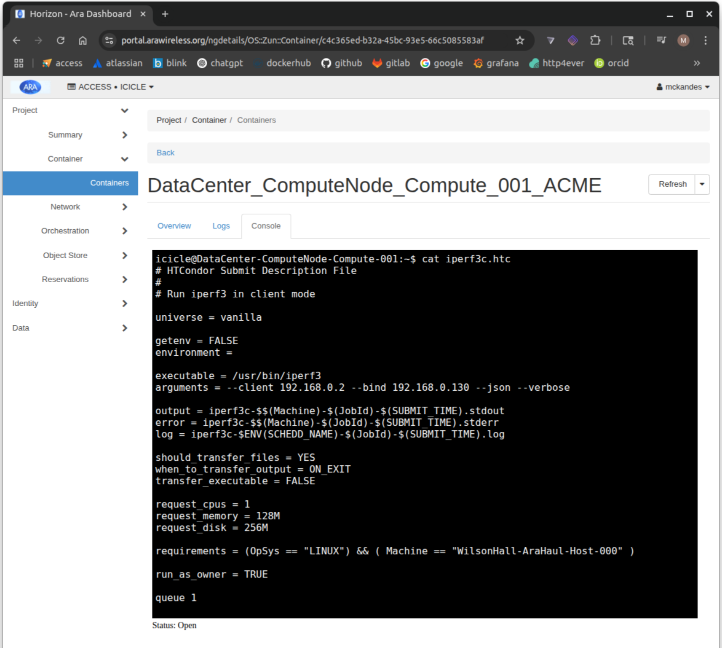
AraHaul Demonstration (with HTCondor)



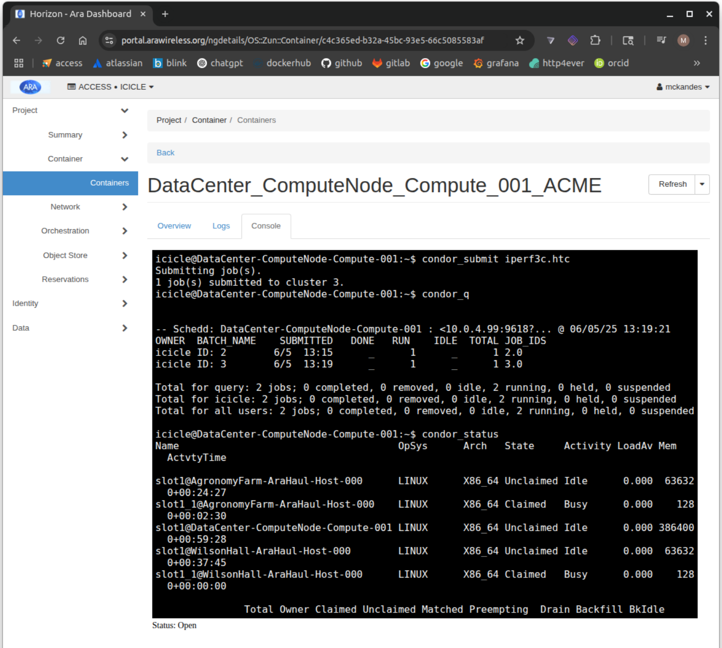
AraHaul Demonstration (with HTCondor)



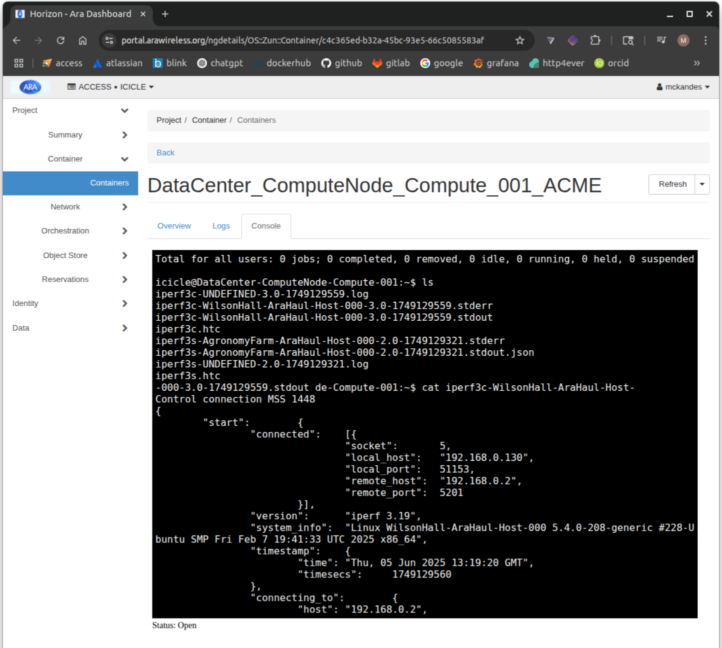
AraHaul Demonstration (with HTCondor)



AraHaul Demonstration (with HTCondor)



AraHaul Demonstration (with HTCondor)



Acknowledgements

https://icicle.ai

This work is funded by the Office of Advanced Cyberinfrastructure (OAC) in the Directorate for Computer and Information Science and Engineering (CISE) at the U.S. National Science Foundation (NSF) as part of the National AI Research Institute for **I**ntelligent **C**yber**I**nfrastructure with **C**omputational **L**earning in the **E**nvironment (ICICLE) under OAC-2112606.

Acknowledgements

ARA

NSF

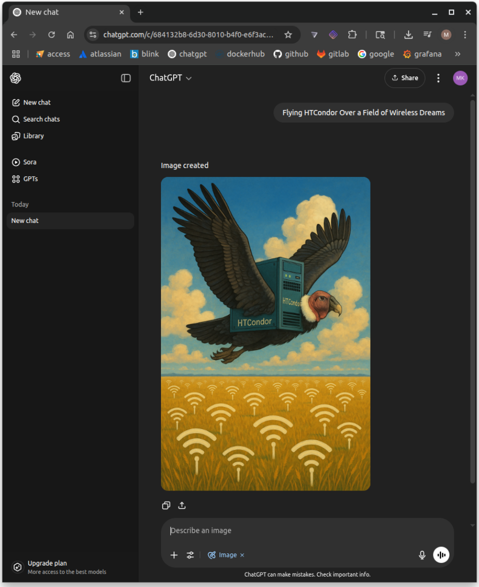
Platforms for Advanced Wireless Research

USDA National Institute of Food and Agriculture
U.S. DEPARTMENT OF AGRICULTURE

ARA is funded by the Division of Computer and Network Systems (CNS) in the Directorate for Computer and Information Science and Engineering (CISE) at the U.S. National Science Foundation (NSF) through its Platforms for Advanced Wireless Research (PAWR) program under CNS-2130889.

Additional funding and in-kind support for ARA has been provided by its PAWR industry partners and the National Institute of Food and Agriculture (NIFA) at the U.S. Department of Agriculture (USDA) through its Agriculture and Food Research Initiative (AFRI) under NIFA-2021-67021-33775.

Questions?


New chat
chatgpt.com/c/684132b8-6d30-8010-b4f0-e6f3ac...
access
atlassian
blink
chatgpt
dockerhub
github
gitlab
google
grafana
ChatGPT
Share
New chat
Search chats
Library
Sora
GPTs
Today
New chat
Flying HTCondor Over a Field of Wireless Dreams
Image created
HTCondor
Describe an image
Image
Upgrade plan
More access to the best models
ChatGPT can make mistakes. Check important info.