www. chameleoncloud.org

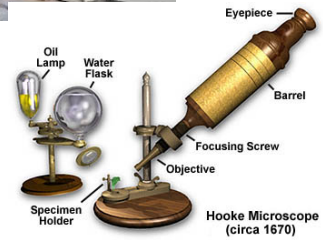
EXPERIMENTS IN THE EDGE TO CLOUD CONTINUUM

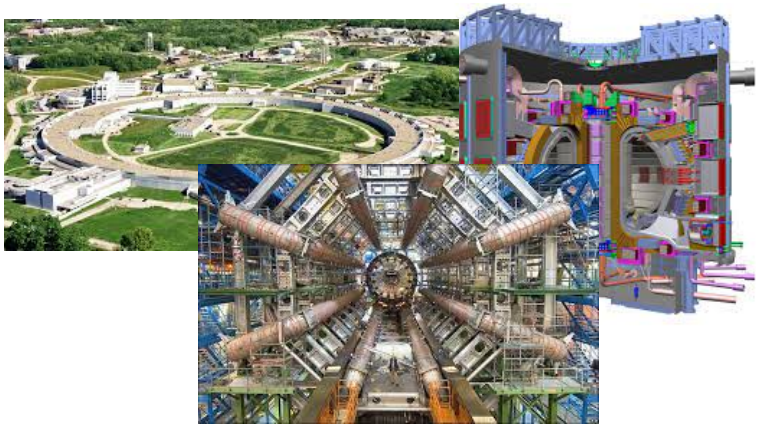
Kate Keahey, Jason Anderson

{keahey, jasonanderson}@uchicago.edu

October 26th, 2021
theNetworkingChannel

SCIENTIFIC INSTRUMENTS





What scientific instruments do Computer Scientists need?

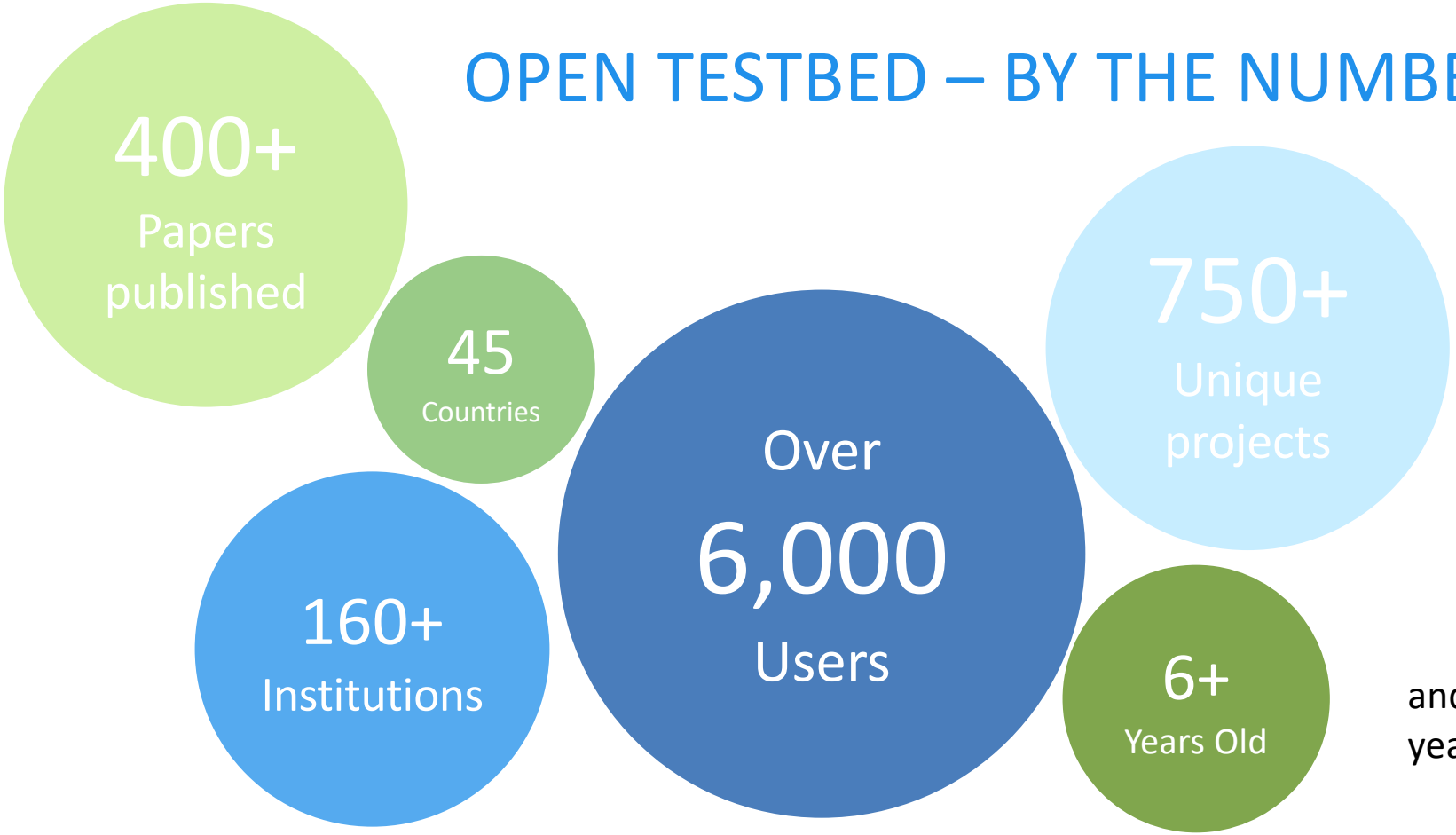
Innovative and diverse hardware, breadth of deployment, freedom to touch and measure every aspect of configuration and behavior.

Constantly evolving!

CHAMELEON IN A NUTSHELL

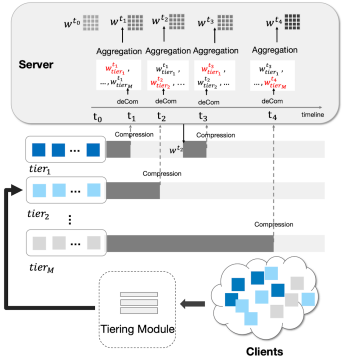
- Chameleons like to change: a testbed that adapts itself to your experimental needs
  - **Deep reconfigurability** (bare metal) and isolation + KVM cloud (different cost/isolation trade-off)
  - Capabilities: power on/off, reboot, custom kernel, serial console access, etc.
- Balance: **large-scale** versus **diverse** hardware
  - Large-scale: ~large homogenous partition (~15,000 cores), ~6 PB of storage originally distributed over 2 sites (**UC, TACC**) connected with 100G network
  - Diverse: ARMs, Atoms, FPGAs, GPUs, Corsa switches, etc.
  - **CHI-in-a-Box** sites at Northwestern, in preparation: NCAR, IIT, and other places
- Cloud++: CHameleon Infrastructure (CHI) via mainstream cloud tech
  - Powered by **OpenStack** with bare metal reconfiguration (Ironic) + "special sauce" (50/50 split)
  - Blazar contribution recognized as official OpenStack component
- Reproducibility, repeatability, and sharing
  - **Jupyter integration** for imperative and non-transactional experiment packaging, **Chameleon daypass** for easy access, **Trovi** for sharing and finding experiments, integration with **Zenodo** for publishing

OPEN TESTBED – BY THE NUMBERS

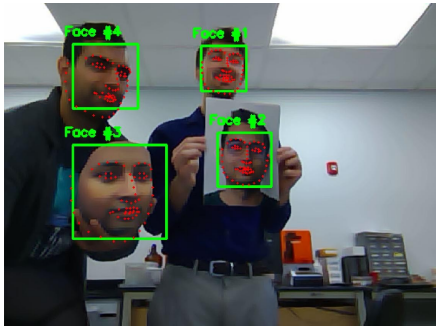


and 3 more years to grow!

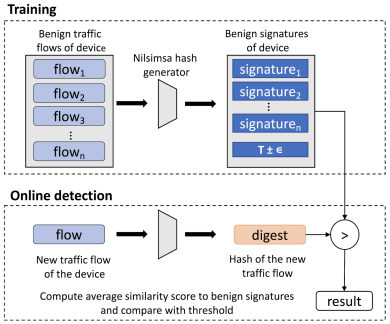
FROM CLOUD TO EDGE WITH CHAMELEON



federated learning



biometrics



network traffic fingerprinting for IoT devices

- Increasingly more Chameleon project applications working on IoT/edge
  - https://chameleoncloud.org/blog/category/user-experiments/
- Simulation/emulation don't always provide the answer: What are the impacts of this approach on power management on edge device? How will the performance transfer to edge? Can we measure the impact of distribution/networking for edge/cloud applications?
- Goal: "realistic edge to cloud experiments from one Jupyter notebook"

WHAT DOES AN EDGE TESTBED LOOK LIKE?

A lot like a cloud!
All the features we know
and love but for edge!

Not at all like a cloud!
Location, location, location!
Not server-class!
IoT: cameras, actuators, SDRs!
And many other challenges!

- CHI@Edge: all the features you know and love plus
  - **Non-prescriptive** access/reconfiguration: container deployment
  - Support for **peripherals** based on an extensible plug-in model
  - **Mixed ownership** model via an SDK with devices, virtual site, and **restricted sharing**
  - Chameleon@Edge Workshop 09/21 https://chameleoncloud.org/chiedge-community-workshop/
  - Edge mailing list: https://groups.google.com/g/chameleon-edge-users

WHAT DOES AN EDGE TESTBED LOOK LIKE?

In-network processing

Network/compute heterogeneity

Network Function Virtualization

Network slicing

Intelligent edge algorithms

Edge to cloud workflows

IoT and wireless multi-tenancy

Latency-aware job placement

Data management for edge

Power management

Operating system for edge

Edge security and privacy

Reliability and Availability

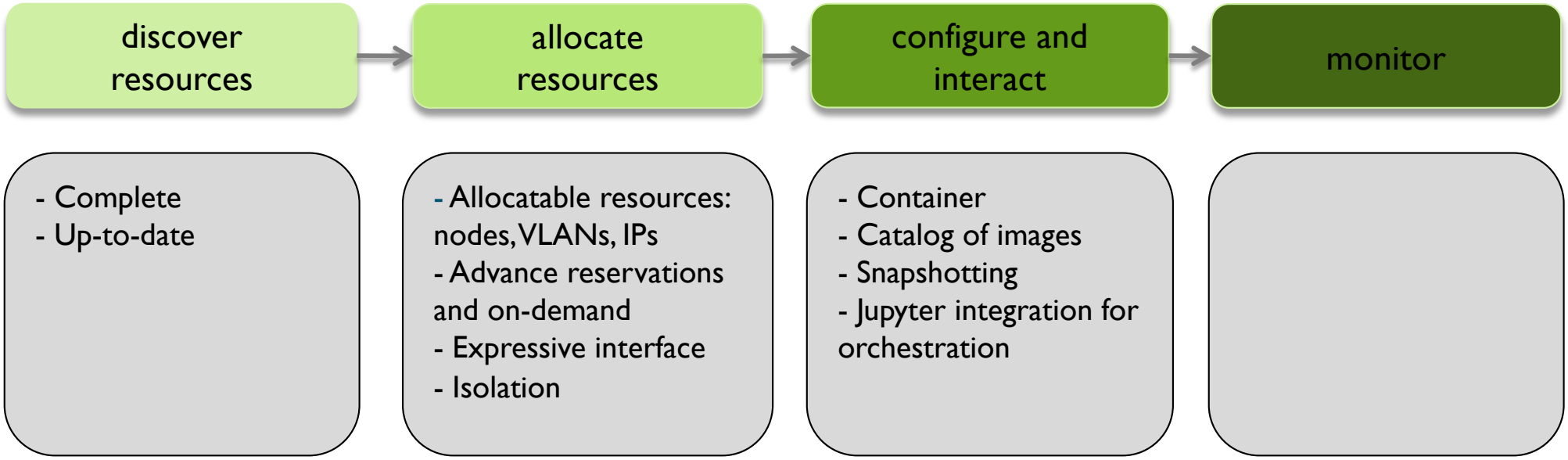
CHI@Edge

chameleon-owned devices

user-owned devices

BUILDING CHI@EDGE

CHI@EDGE EXPERIMENTAL WORKFLOW (PREVIEW)

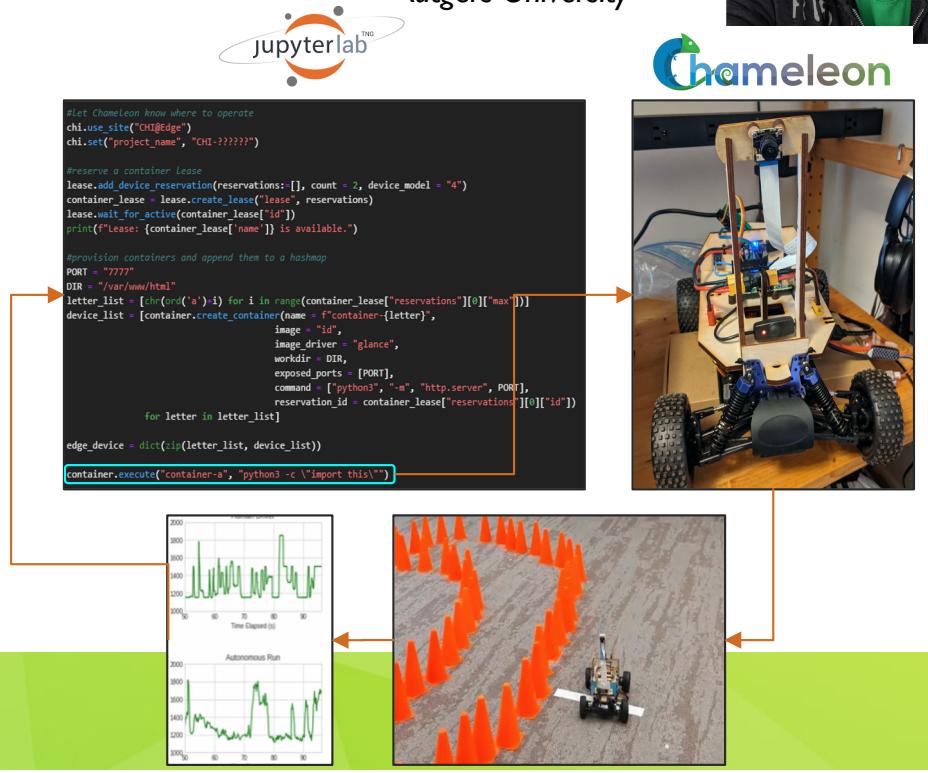


Authentication via federated identity, accessed via GUI, CLI and python/Jupyter

AUTONOMOUS CARS WITH CHI@EDGE

- Goal:
  - Teach machine learning and systems concepts using remote autonomous cars
- Challenges:
  - Control the cars remotely: manual workflows require lots of teacher effort
  - Iterate on code while learning and exploring
  - Collect, store, and process large datasets
- CHI@Edge:
  - Car reservations
  - Access through JupyterHub
  - Provides consistent network connection
  - Deploy code and collect results with repeatable workflows

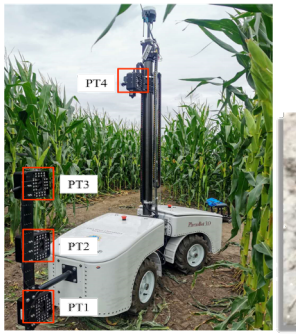
Rick Anderson
Virtual Worlds, Director
Rutgers University



ARA: WIRELESS LIVING LAB FOR SMART & CONNECTED RURAL COMMUNITIES

Hongwei Zhang, ARA PI
Iowa State University

- ARA objectives
  - Enable research to achieve a factor of 10+ reduction in broadband cost and make rural broadband as affordable as urban broadband!
  - Support broadband use cases for rural communities
- ARA wireless living lab
  - Deploy advanced wireless platforms in Central Iowa (>600 square miles); capture systems and application and community contexts of rural broadband
  - Mainstream open-source platforms for living lab management and experimentation: OpenStack, CHI-in-a-Box & CHI@Edge, ONF (SD-RAN, SD-CORE, ONOS), srsRAN, OpenAirInterface etc
  - CHI@Edge: collaborating on spectrum reservations for management of wireless networks and CHI@Edge in a Box

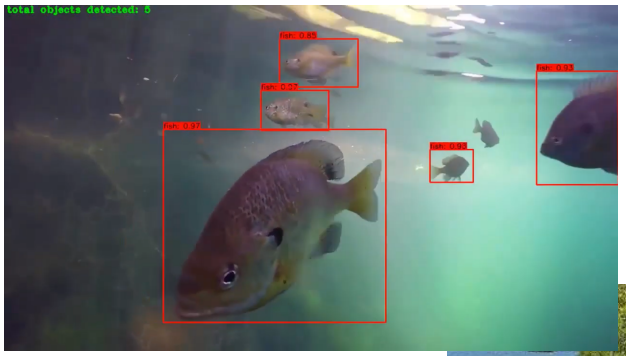


Location and Interior view of ISU Beef Nutrition Research Farm

EDGE FOR MARINE BIOLOGY

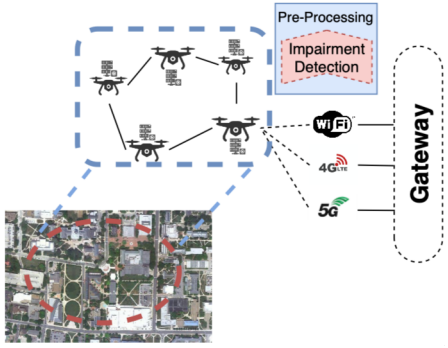
- Goal: map existing fish populations and thereby understand better how pollution impacts their habitat and the general Biscayne Bay ecosystem
- Challenges: What is the best cloud/edge strategy for collecting and analyzing data from the autonomous vehicle (AV)? How does the resolution of video data and quality of network connection influence them?
- CHI@Edge: using CHI@Edge for developing edge to cloud data processing workflows via Jupyter notebooks

Kevin Boswell, Leonardo Bobadilla, and Jonathan Tsen
Florida International University

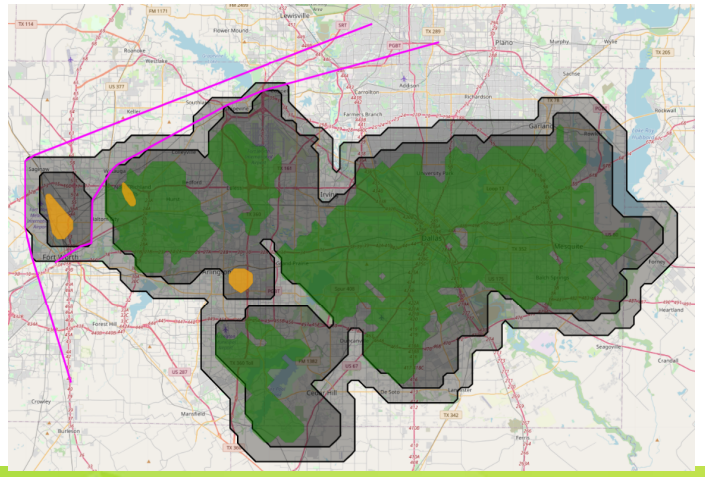


FLYNET: AN 'ON-THE-FLY' PROGRAMMABLE END-TO-END NETWORK-CENTRIC PLATFORM

- Architecture and tools that support edge computing devices in scientific workflows
- Critical for low latency and ultra-low latency applications: e.g., drone video analytics and route planning for drones
- Challenges: integration of compute and networking infrastructure, in-network processing, end-to-end monitoring, workflow management (Pegasus)
- CHI@Edge
  - Use for edge computing experiments
  - Provide experiments that can be reproduced by other researchers
  - FlyNet to provide tools to allow researcher to include CHI@Edge in their workflows



Mike Zink FlyNet PI
U of Mass, Amherst



PARTING THOUGHTS

- Constantly in motion: scientific instruments are laying down the pavement as science walks on it
- **Testbed inversion**: from cloud to edge
  - Before: expensive **provider-owned** hardware as the main draw
  - Now: **user-owned** inexpensive hardware using testbed **sharing and connecting** mechanisms
  - Testbeds == effective **sharing and connecting** mechanism + residual resources
- **Heterogeneity** of resource sharing: deep reconfiguration has a cost
  - From one-size-fit-all: bare metal to KVM to containers and beyond
- Sharing research in digital form: we can no longer afford not to

Think Big!

(with the help of a small reptile)

www.chameleoncloud.org