



www.chameleoncloud.org

CHAMELEON: REPRODUCIBLE EXPERIMENTS IN THE EDGE TO CLOUD CONTINUUM

Kate Keahey

Mathematics and CS Division, Argonne National Laboratory

CASE, University of Chicago

keahey@anl.gov

January 7th, 2022

IEEE FNI Testbed Workshop



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: diversity and scale – from large to small
 - ▶ Large to small: from 2 supercomputing sites (**UC, TACC**) connected with 100G network to **edge devices**
 - ▶ Diverse: FPGAs, GPUs, NVMe, NVDIMMs, Corsair switches, edge devices via CHI@Edge 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, **Trove** for sharing and finding experiments, integration with **Zenodo** for publishing



OPEN TESTBED – BY THE NUMBERS

450+

Papers
published

45

Countries

800+

Unique
projects

6,000+

Users

160+

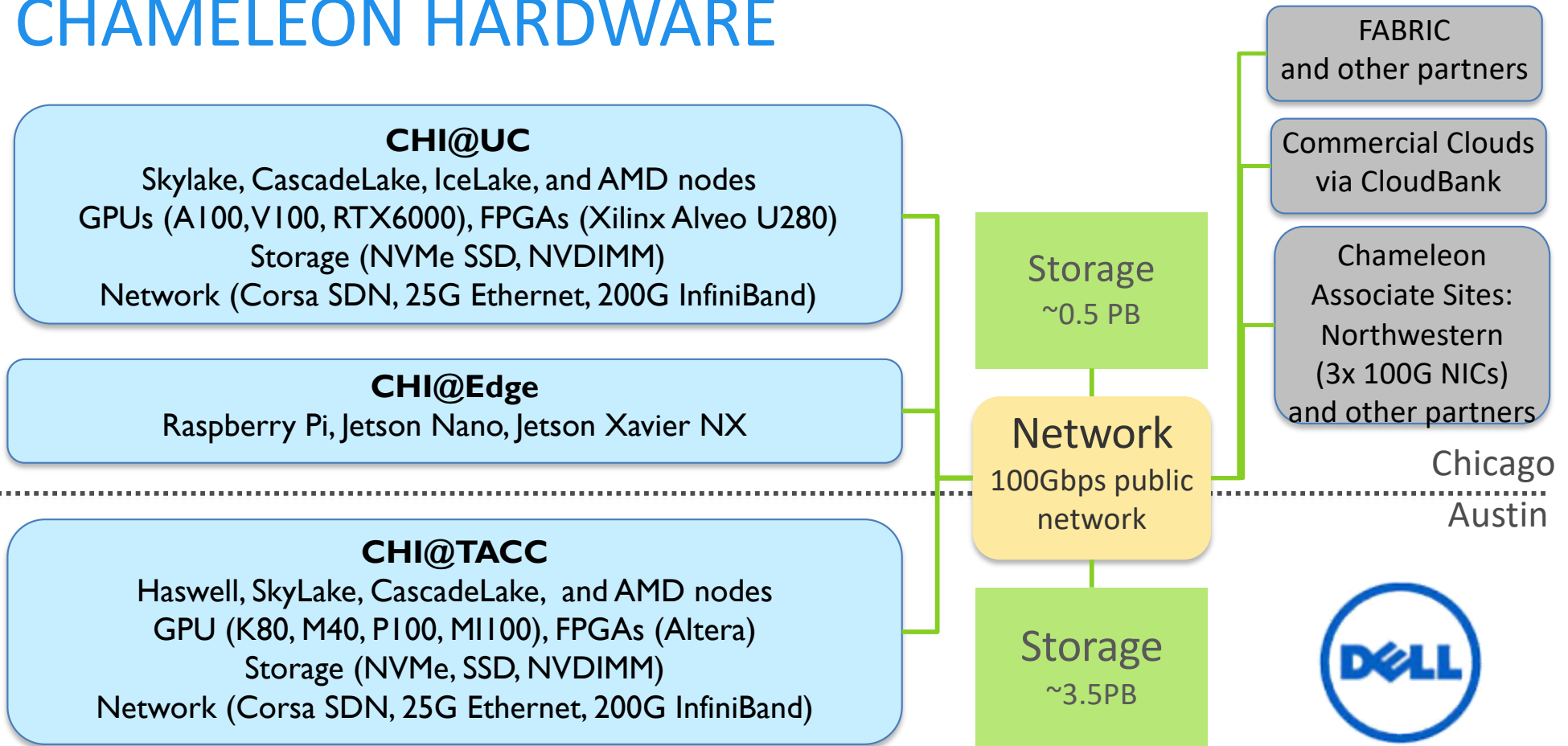
Institutions

6+

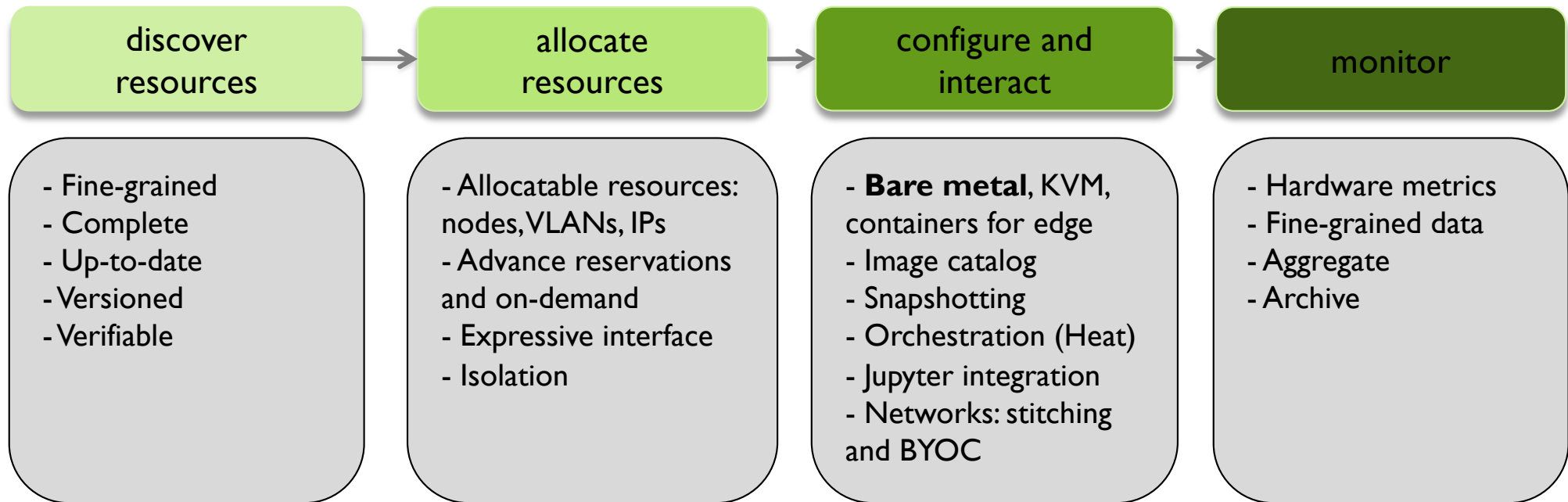
Years Old

~3 more years
to grow!

CHAMELEON HARDWARE



CHI EXPERIMENTAL WORKFLOW



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

Paper: "Lessons Learned from the Chameleon Testbed", USENIX ATC 2020

NEW IN P3: CHI@EDGE (PREVIEW)



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!

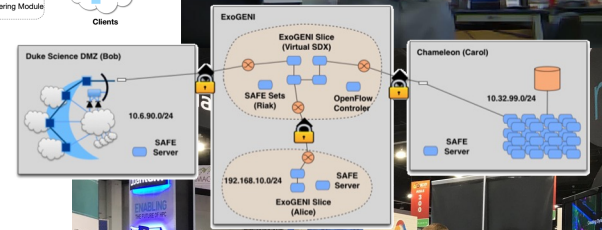
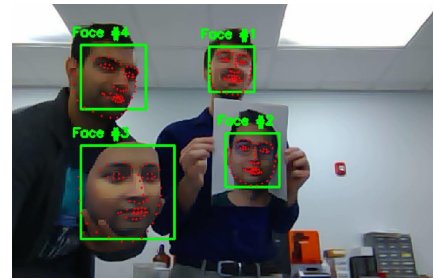
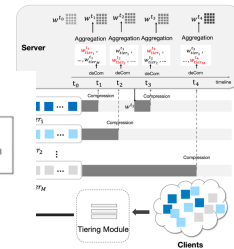
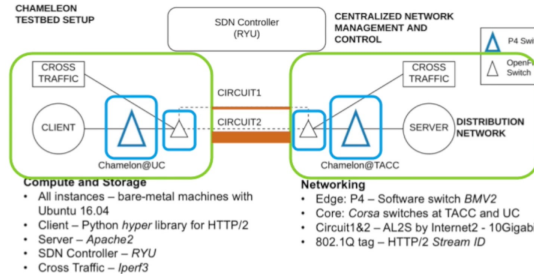
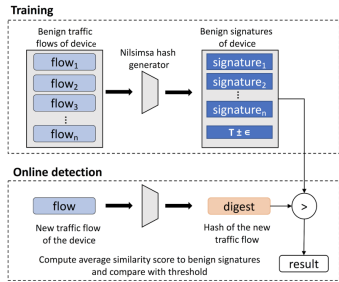
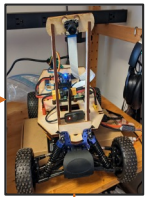


- ▶ CHI@Edge: all the features you know and love plus
 - ▶ Reconfiguration via container deployment
 - ▶ Support for peripherals based on an extensible plug-in model
 - ▶ **Mixed ownership** model via an SDK with devices, virtual site, and **restricted sharing**
 - ▶ And more... Chameleon@Edge Community Workshop in September 2021 <https://chameleoncloud.org/chiedge-community-workshop/>

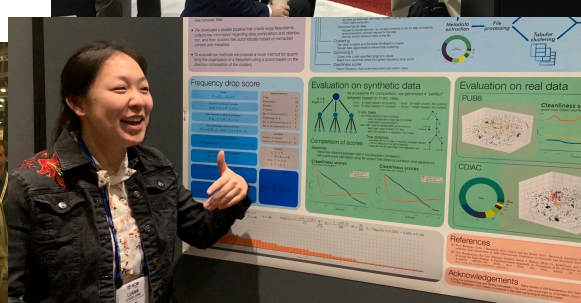
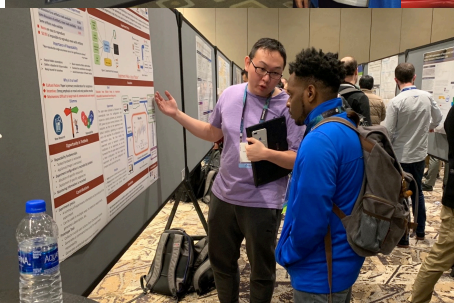
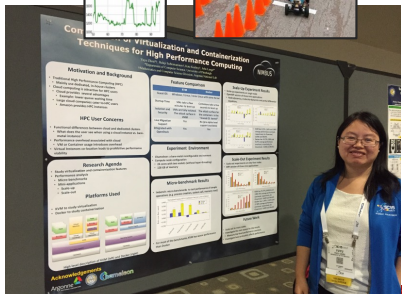


LEAVING NO EXPERIMENT BEHIND!

```
#!/usr/bin/perl use strict; use warnings; my $device = "10.10.10.1"; my $signature = "10.10.10.1"; my $signature2 = "10.10.10.1"; my $signature3 = "10.10.10.1"; my $signature4 = "10.10.10.1"; my $signature5 = "10.10.10.1"; my $signature6 = "10.10.10.1"; my $signature7 = "10.10.10.1"; my $signature8 = "10.10.10.1"; my $signature9 = "10.10.10.1"; my $signature10 = "10.10.10.1"; my $signature11 = "10.10.10.1"; my $signature12 = "10.10.10.1"; my $signature13 = "10.10.10.1"; my $signature14 = "10.10.10.1"; my $signature15 = "10.10.10.1"; my $signature16 = "10.10.10.1"; my $signature17 = "10.10.10.1"; my $signature18 = "10.10.10.1"; my $signature19 = "10.10.10.1"; my $signature20 = "10.10.10.1"; my $signature21 = "10.10.10.1"; my $signature22 = "10.10.10.1"; my $signature23 = "10.10.10.1"; my $signature24 = "10.10.10.1"; my $signature25 = "10.10.10.1"; my $signature26 = "10.10.10.1"; my $signature27 = "10.10.10.1"; my $signature28 = "10.10.10.1"; my $signature29 = "10.10.10.1"; my $signature30 = "10.10.10.1"; my $signature31 = "10.10.10.1"; my $signature32 = "10.10.10.1"; my $signature33 = "10.10.10.1"; my $signature34 = "10.10.10.1"; my $signature35 = "10.10.10.1"; my $signature36 = "10.10.10.1"; my $signature37 = "10.10.10.1"; my $signature38 = "10.10.10.1"; my $signature39 = "10.10.10.1"; my $signature40 = "10.10.10.1"; my $signature41 = "10.10.10.1"; my $signature42 = "10.10.10.1"; my $signature43 = "10.10.10.1"; my $signature44 = "10.10.10.1"; my $signature45 = "10.10.10.1"; my $signature46 = "10.10.10.1"; my $signature47 = "10.10.10.1"; my $signature48 = "10.10.10.1"; my $signature49 = "10.10.10.1"; my $signature50 = "10.10.10.1"; my $signature51 = "10.10.10.1"; my $signature52 = "10.10.10.1"; my $signature53 = "10.10.10.1"; my $signature54 = "10.10.10.1"; my $signature55 = "10.10.10.1"; my $signature56 = "10.10.10.1"; my $signature57 = "10.10.10.1"; my $signature58 = "10.10.10.1"; my $signature59 = "10.10.10.1"; my $signature60 = "10.10.10.1"; my $signature61 = "10.10.10.1"; my $signature62 = "10.10.10.1"; my $signature63 = "10.10.10.1"; my $signature64 = "10.10.10.1"; my $signature65 = "10.10.10.1"; my $signature66 = "10.10.10.1"; my $signature67 = "10.10.10.1"; my $signature68 = "10.10.10.1"; my $signature69 = "10.10.10.1"; my $signature70 = "10.10.10.1"; my $signature71 = "10.10.10.1"; my $signature72 = "10.10.10.1"; my $signature73 = "10.10.10.1"; my $signature74 = "10.10.10.1"; my $signature75 = "10.10.10.1"; my $signature76 = "10.10.10.1"; my $signature77 = "10.10.10.1"; my $signature78 = "10.10.10.1"; my $signature79 = "10.10.10.1"; my $signature80 = "10.10.10.1"; my $signature81 = "10.10.10.1"; my $signature82 = "10.10.10.1"; my $signature83 = "10.10.10.1"; my $signature84 = "10.10.10.1"; my $signature85 = "10.10.10.1"; my $signature86 = "10.10.10.1"; my $signature87 = "10.10.10.1"; my $signature88 = "10.10.10.1"; my $signature89 = "10.10.10.1"; my $signature90 = "10.10.10.1"; my $signature91 = "10.10.10.1"; my $signature92 = "10.10.10.1"; my $signature93 = "10.10.10.1"; my $signature94 = "10.10.10.1"; my $signature95 = "10.10.10.1"; my $signature96 = "10.10.10.1"; my $signature97 = "10.10.10.1"; my $signature98 = "10.10.10.1"; my $signature99 = "10.10.10.1"; my $signature100 = "10.10.10.1";
```



Supporting research projects in architecture, operating systems design, virtualization, power management, real-time analysis, security, storage systems, databases, networking, machine learning, neural networks, data science, and many others.



ARA: WIRELESS LIVING LAB FOR SMART & CONNECTED RURAL COMMUNITIES

▶ ARA objectives

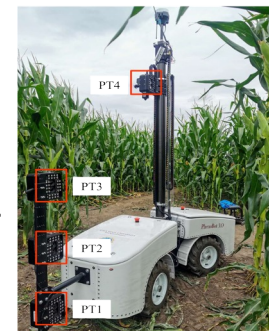
- ▶ Enable research to achieve a factor of 10+ reduction in broadband cost and make rural broadband as affordable as urban broadband!
- ▶ Support broadband use cases for rural communities

▶ 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



Hongwei Zhang, ARA PI
Iowa State University

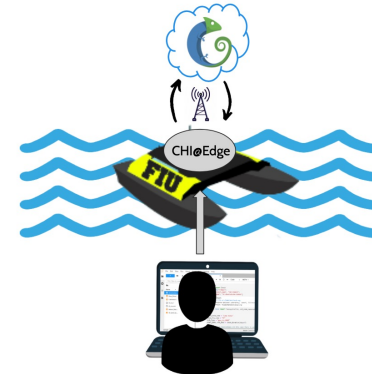
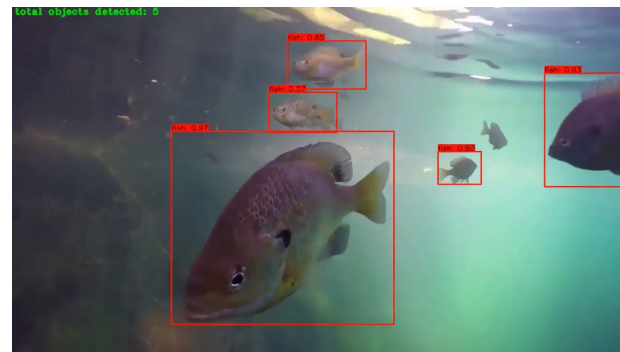
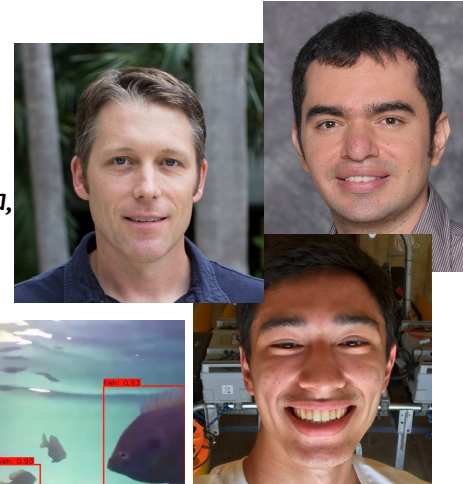


Location and Interior view of
ISU Beef Nutrition Research Farm

EDGE FOR MARINE BIOLOGY

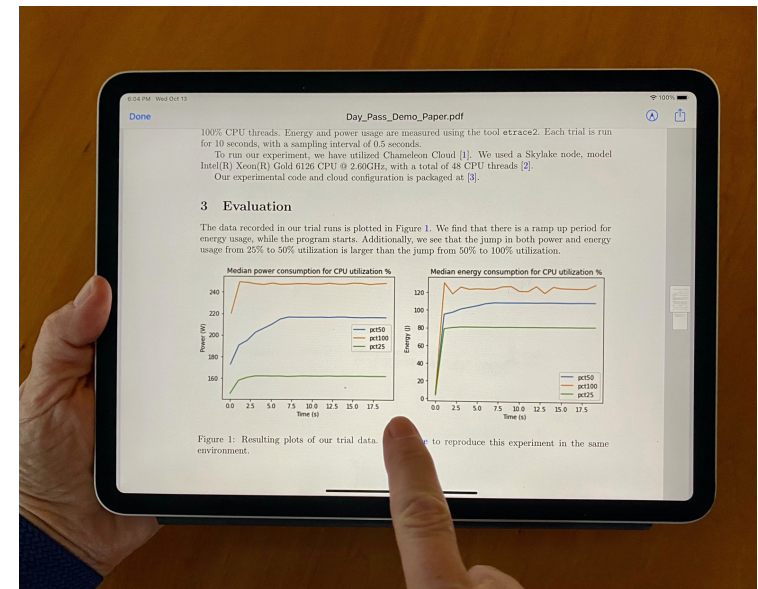
- ▶ Goal: map existing fish populations and thereby understand better how pollution impacts their habitat and the general Biscayne Bay ecosystem
- ▶ Challenges: What is the best cloud/edge strategy for collecting and analyzing data from the autonomous vehicle (AV)? How does the resolution of video data 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



PRACTICAL REPRODUCIBILITY

- ▶ Can experiments be as sharable as papers are today?
- ▶ Could it be as easy to provide conditions for reviewers to repeat experiments or data analysis in a paper as it is to organize a PC meeting?
- ▶ Can I simply integrate somebody's model into my research instead of reinventing the wheel?
- ▶ Can I have so much fun playing with somebody's experiment that discover a new result?
- ▶ Can I develop exercises for my class based on most recent research results?



The existence of powerful open testbeds is a fundamental requirement for practical reproducibility

TESTBED AS SHARING PLATFORM

- ▶ **Instruments held in common** are a reproducibility imperative
 - ▶ Hardware and hardware versions: >105 versions over 5 years
 - ▶ Expressive allocation
- ▶ **Sharing via **cloud pattern****
 - ▶ Disk images, orchestration templates, and other artifacts
 - ▶ Chameleon >130,000 images, >35,000 orchestration templates and counting
- ▶ Testbed as “player” for environments



Paper: “The Silver Lining”, IEEE Internet Computing 2020

WHAT IS MISSING?

- ▶ Packaging: complete, imperative, non-transactional, integrated (literate programming)
- ▶ Get access for reproducibility
- ▶ Discover/find experiments through various channels

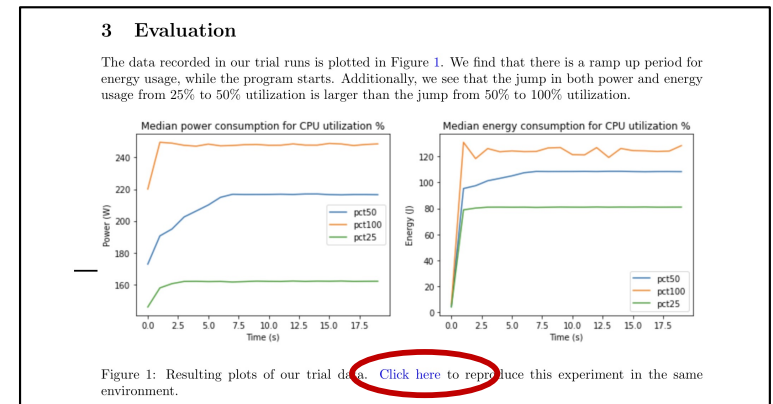


- ▶ Package experiment in a way that is cost-effective but also user-friendly
- ▶ Give access for reproducibility
- ▶ Share work in progress; publish and advertise completed work



TESTBED ACCESS WITH CHAMELEON DAYPASS

- ▶ Authors create a subproject with multiple short-term leases that are long enough to reproduce the experiment
- ▶ Readers click through data of a published experiment, request a daypass, and reproduce either the experiment or data analysis



Chameleon About Learn Experiment Blog Log in

Artifacts / Getting Started with Chameleon: Power management experiment

Getting Started with Chameleon: Power management experiment

This notebook is a short example of how to use Chameleon notebooks to run a simple experiment, and analyze the data, using the python-chi interface.

Estimated duration: 1 hour
Support contact: help@chameleoncloud.org

2 Sept. 29, 2021, 12:44 p.m. example experiment

Authors
Jason Anderson (University of Chicago)
Mark Powers (University of Chicago)

Request day pass

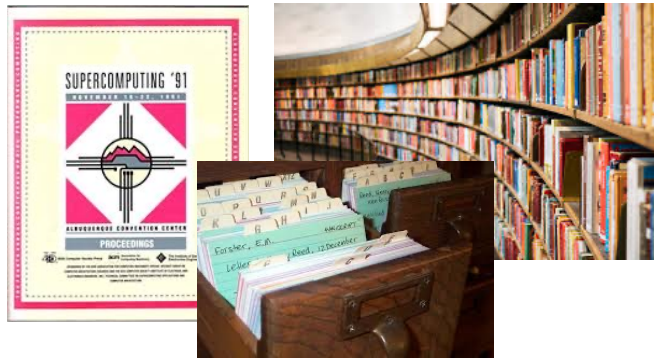
If you do not have an active Chameleon allocation, or would prefer to not use your allocation, you can request a temporary one from the PI of the project this artifact belongs to.

Versions

Version 2	Sept. 29, 2021, 12:43 p.m.
Version 1	Sept. 29, 2021, 12:37 p.m.

SHARING AND FINDING EXPERIMENTS

Familiar research sharing ecosystem



Digital research sharing ecosystem



- ▶ Digital publishing with Zenodo: make your experiments citable via Digital Object Identifiers (DOIs)
- ▶ Trovi: sharing work in progress
 - ▶ BINs to collect all the artifacts, fine-grained sharing, versioning
 - ▶ Portal to browse, filter, and find interesting experiments
 - ▶ Integrated with Jupyter/Chameleon, Swift, Zenodo, and github (in progress)



PARTING THOUGHTS

- ▶ Constantly in motion: scientific instruments are laying down the pavement as science walks on it
- ▶ **Testbed evolution:** 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
- ▶ **Sharing your research digitally is more important than ever!**
 - ▶ Make it easy with Chameleon: public platform, environments as images, packaging, access, and sharing mechanisms at the ready
 - ▶ Biggest benefit in emergent area == real incentives

Think Big!

(with the help of a small reptile)



www.chameleoncloud.org