Enabling P4 Hands-on Training in an Academic Cloud

Jose Gomez, Elie Kfoury, Jorge Crichigno
University of South Carolina

International Workshop on Test and Evaluation of Programmable Networks
Marina Del Rey, Los Angeles, California

June 1, 2022

Agenda

- Motivation for virtual labs and Academic Cloud
- Academic Cloud
- POD design and lab libraries
- Using the Academic Cloud
- Relevant features
- Concluding remarks

Motivation for Virtual Labs and Academic Cloud

- According to the IEEE and ACM[1], the IT curriculum should emphasize "learning IT core concepts with authentic practice" and "use of professional tools and platforms"
  - "It is not enough to simply attend courses and read books. Hands-on learning is essential…"
- Using physical laboratories has been challenging
  - Difficult to scale
  - Expensive (space, maintenance, staff)
  - Since COVID-19 emerged, the capacity of labs has been further reduced (distance requirements)

1. Information Technology Curricula 2017, ACM/IEEE Joint Committee. Online: https://tinyurl.com/4nqqwa5m.

Motivation for Virtual Labs and Academic Cloud

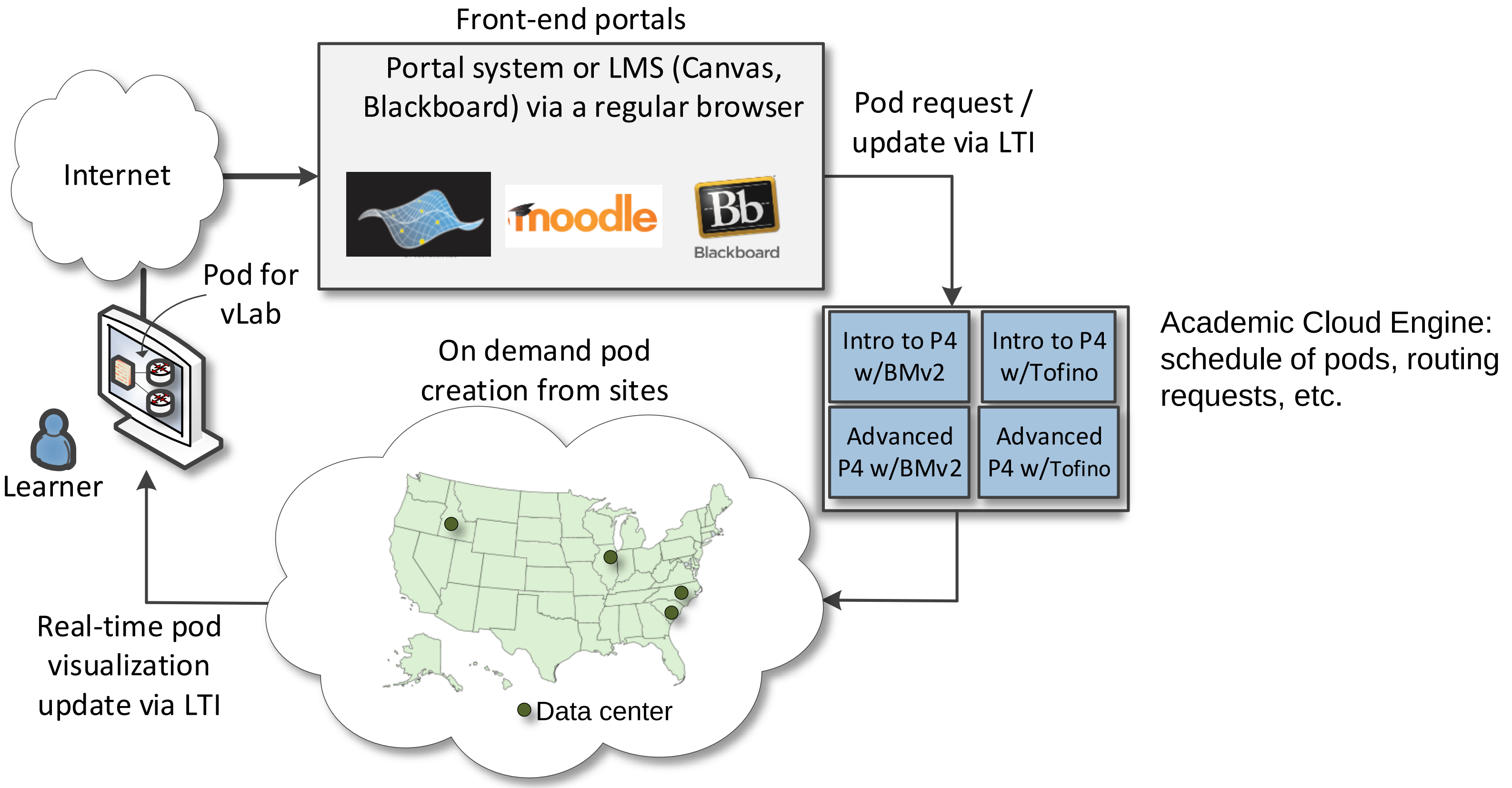
- “The Missing Millions”(NSF report – Oct. 2021. https://tinyurl.com/5awhdazy)
- A report on what can be done to reach out those who are yet to be engaged in STEM workforce
- 15 focus groups, experts on research computing infrastructure
  - “The present research computing and data ecosystems look impenetrable to many of those not yet engaged…”
  - “Lower barriers to entry, but build up the controls at the same time”
  - “Invest in cyberinfrastructure and community laboratories at the edge, enabling broader and more diverse participation in science and engineering”
  - “Explore investments in research computing and data infrastructure approaches that are easily accessible (such as GUIs, science apps, and field tools)”

Academic Cloud

- The University of South Carolina (USC) (SC), the Network Development Group (NDG) (NC), and Stanly Community College (SCC) (NC) are deploying the Academic Cloud
- A system dedicated to teaching, training, and research
- The Academic Cloud provides remote-access capability to lab equipment via Internet
- It seamlessly pools and shares resources (CPU, memory, storage) from four data centers; resources are allocated to run virtual laboratories

Academic Cloud

- Data center locations: USC (South Carolina), SCC (North Carolina), NDG (IL), and Idaho National Laboratory (ID)

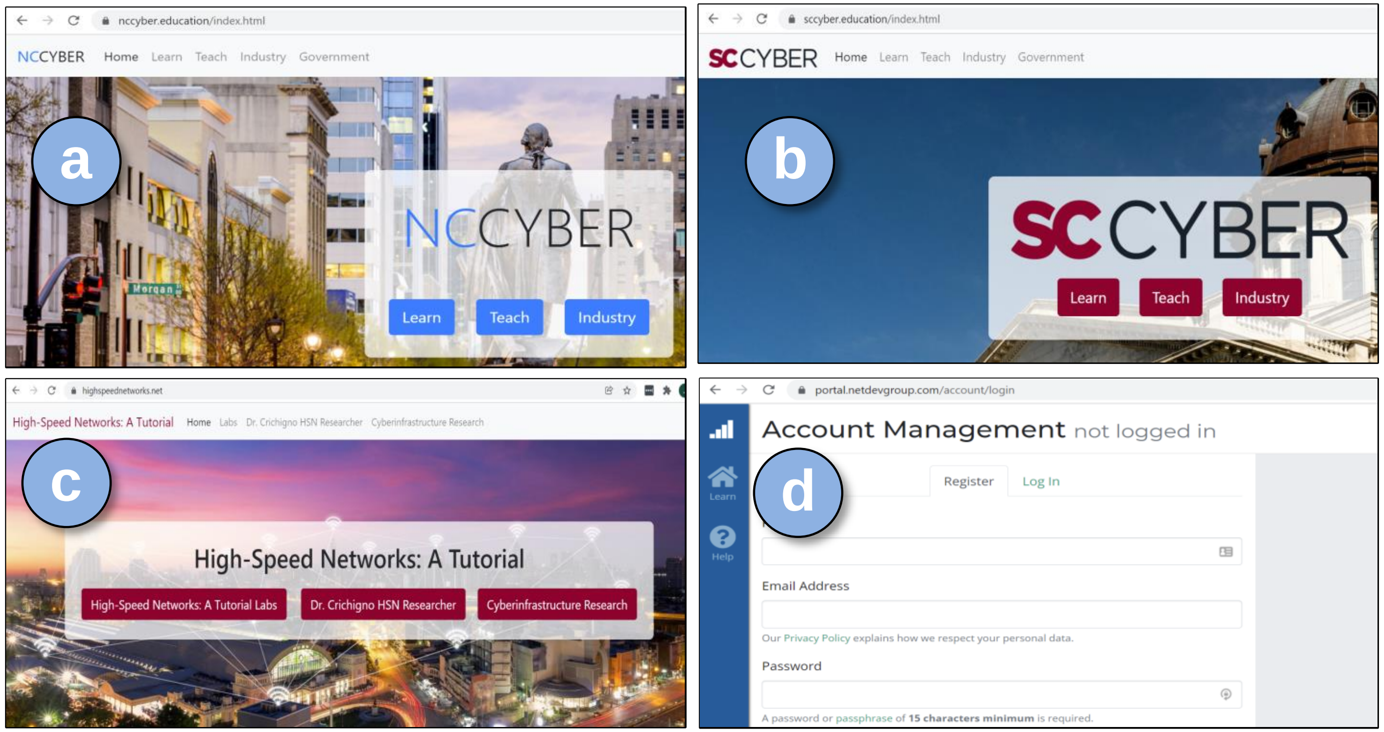


LMS: Learning Management System. LTI: Learning Tools Interoperability

Academic Cloud

- Data center locations: USC (South Carolina), SCC (North Carolina), NDG (IL), and Idaho National Laboratory (ID)

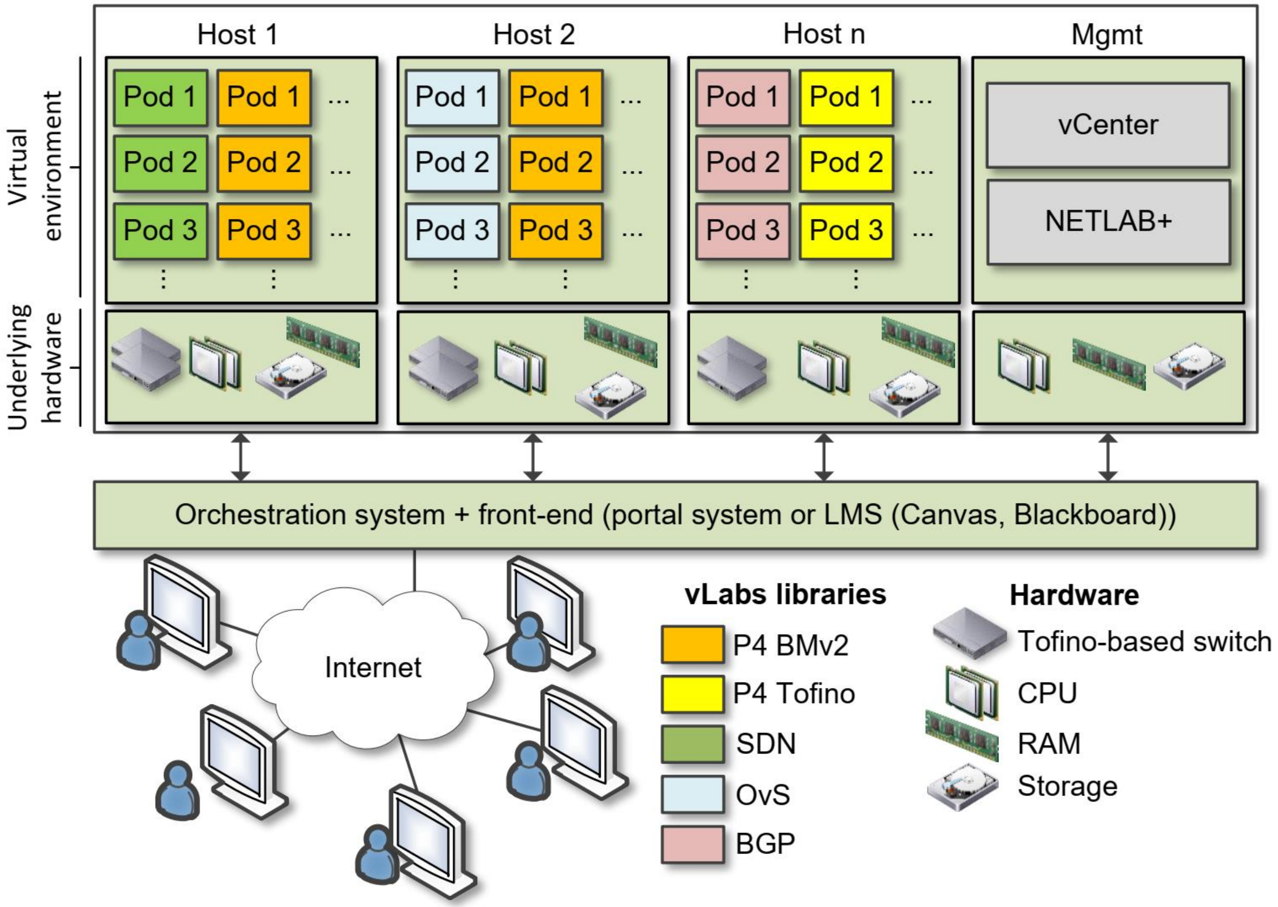
Front-end portals



(a) NC Cyber; (b) SC Cyber; (b) Companion material for a book; (d) General access

Inside a Data Center

- Hosts 1-n store virtual machines (VMs) for virtual labs
- Management server runs vCenter, Management Software (NETLAB+)
- Partnership with Network Development Group (NDG)[1]



1. Network Development Group (NDG). Online: https://netdevgroup.com

Inside a Data Center

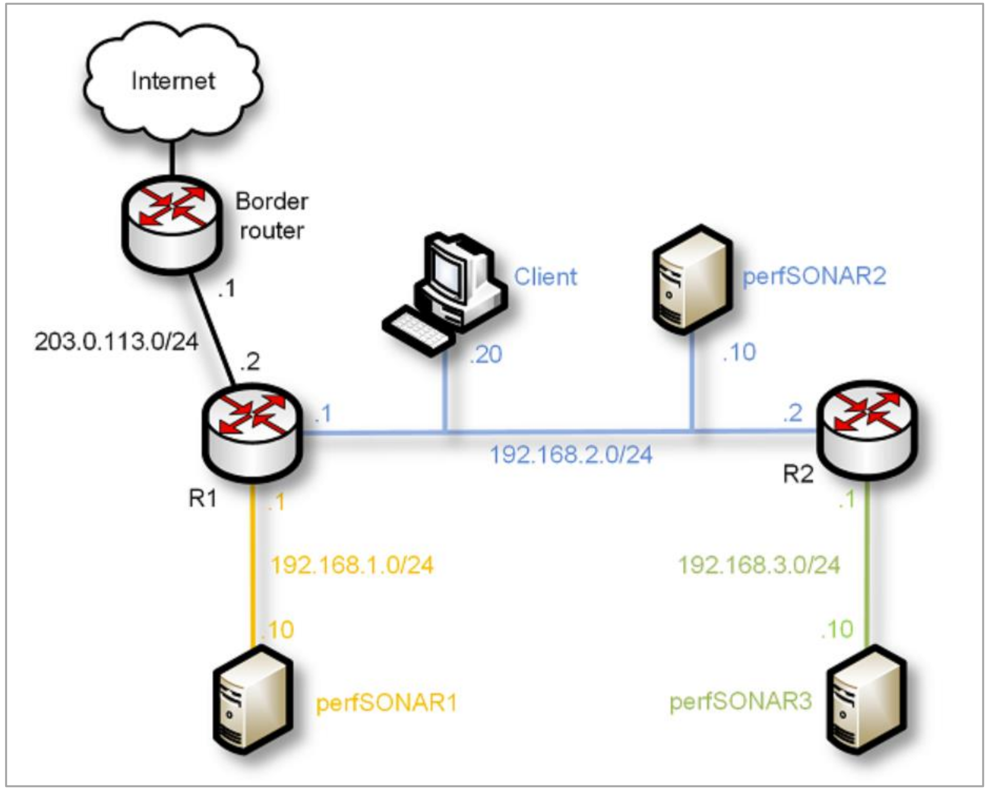
- Example: Stanly Community College

Device	Cores	Storage (TBs)	RAM Memory (GB)
Server 1 (management SCC)	20	12	264
Server 2 (hosting vLabs pods)	32	4	768
Server 3 (hosting vLabs pods)	32	4	768
Server 4 (hosting vLabs pods)	32	4	768
Server 5 (hosting vLabs pods)	32	4	768
Server 6 (hosting vLabs pods)	32	4	768
Server 7 (hosting vLabs pods)	48	1.92	768
Server 8 (hosting vLabs pods)	48	1.92	768
Server 9 (hosting vLabs pods)	48	1.92	768
TOTAL	**324**	**37.76**	**6408**

POD Design

- A virtual laboratory experiment requires a **pod** of devices, or simply pod
- Example: perfSONAR library

POD for perfSONAR labs

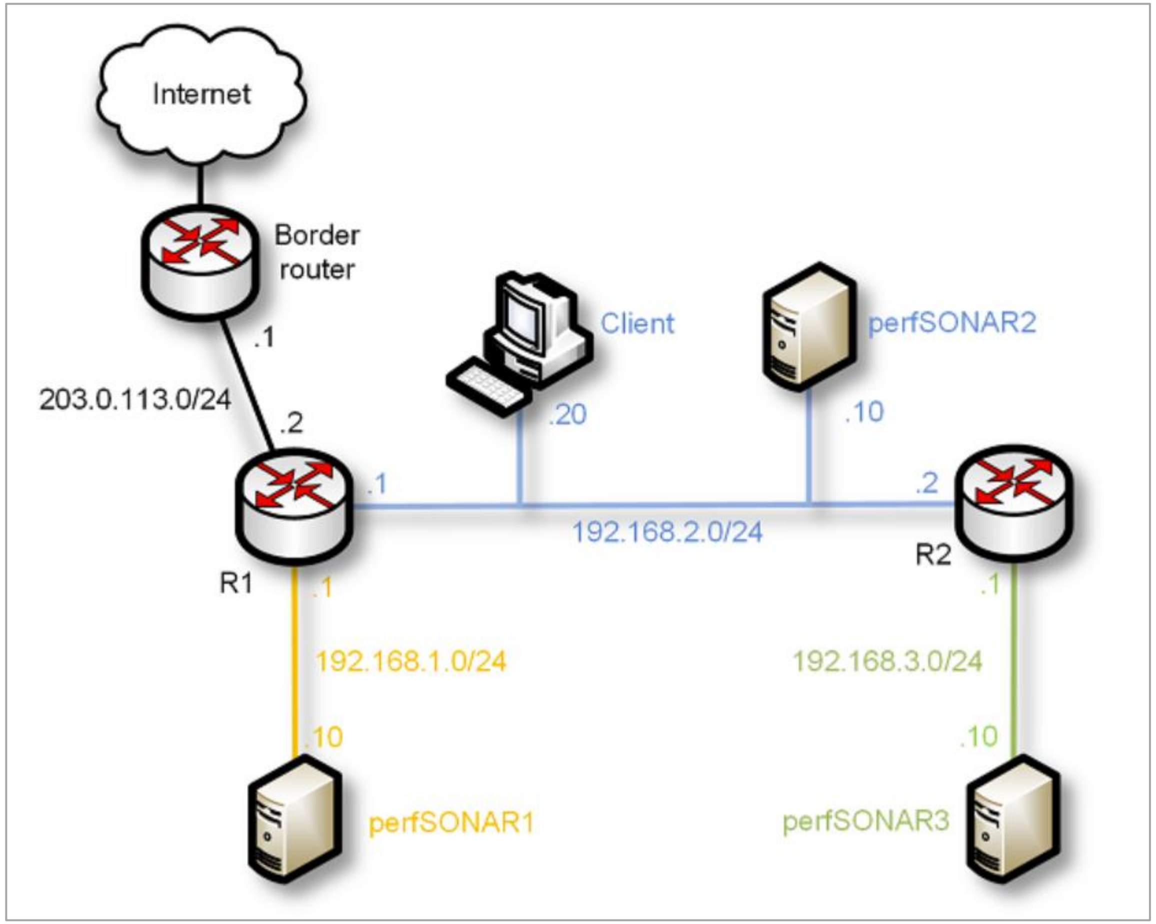


perfSONAR labs

Lab 1	Configuring Administrative Information Using perfSONAR Toolkit GUI
Lab 2	PerfSONAR Metrics and Tools
Lab 3	Configuring Regular Tests Using perfSONAR GUI
Lab 4	Configuring Regular Tests Using pScheduler CLI Part I
Lab 5	Configuring Regular Tests Using pScheduler CLI Part II
Lab 6	Bandwidth-delay Product and TCP Buffer Size
Lab 7	Configuring Regular Tests Using a pSConfig Template
Lab 8	perfSONAR Monitoring and Debugging Dashboard
Lab 9	pSConfig Web Administrator
Lab 10	Configuring pScheduler Limits

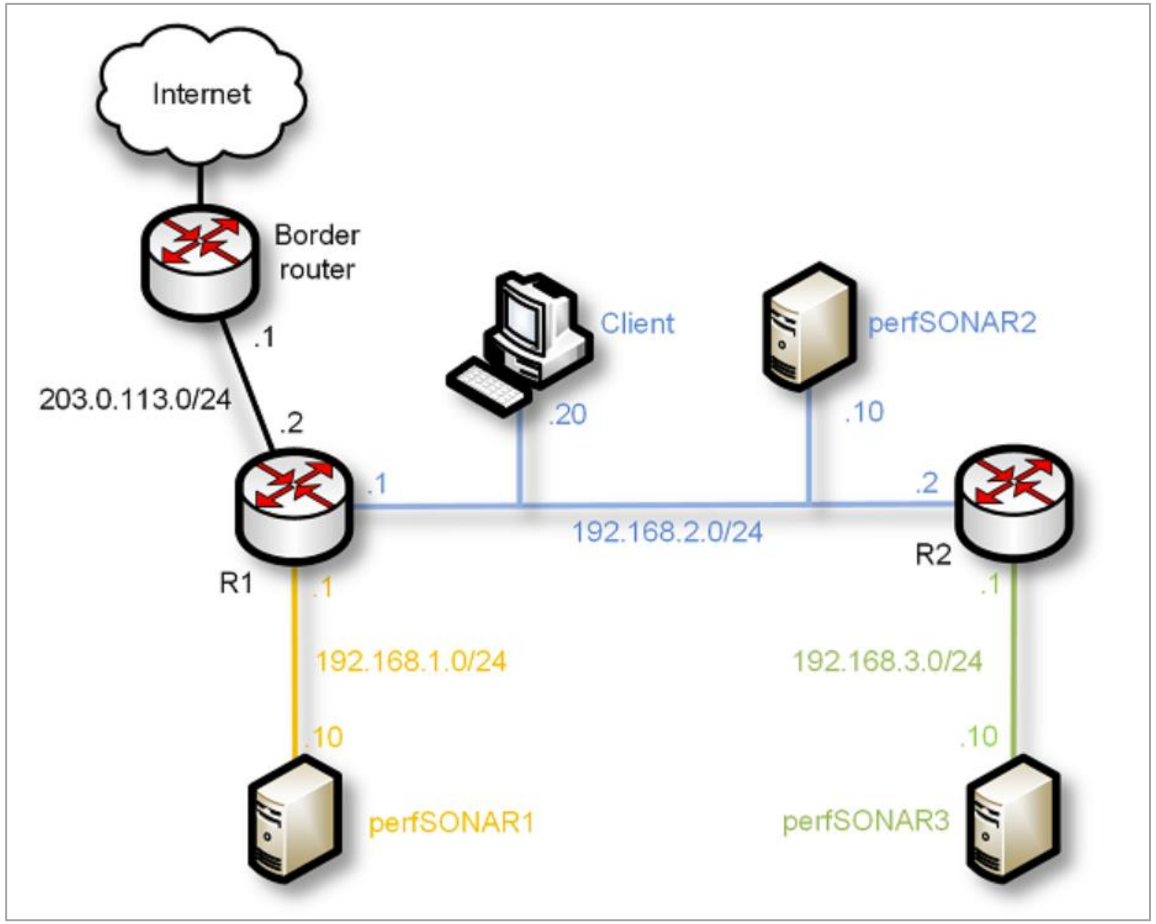
POD Design

- Details of perfSONAR pod
  - Four networks
  - Three servers
  - One client
  - Three routers
  - Connectivity to the Internet
  - Total of seven heterogeneous VMs



POD Design

- Details of perfSONAR pod
  - PODs running simultaneously use the same block of IP addresses
  - Lab manuals are uniform
  - There is a master pod in the system
  - Linked clone VMs are created from the master pod VMs



Introduction to P4 Lab Series

Lab experiments

Lab 1: Introduction to Mininet

Lab 2: Introduction to P4 and BMv2

Lab 3: P4 Program Building Blocks

Lab 4: Parser Implementation

Lab 5: Introduction to Match-action Tables (Part 1)

Lab 6: Introduction to Match-action Tables (Part 2)

Lab 7: Populating and Managing Match-action Tables

Lab 8: Checksum Recalculation and Packet Deparsing

Exercises

Exercise 1: Building a Basic Topology

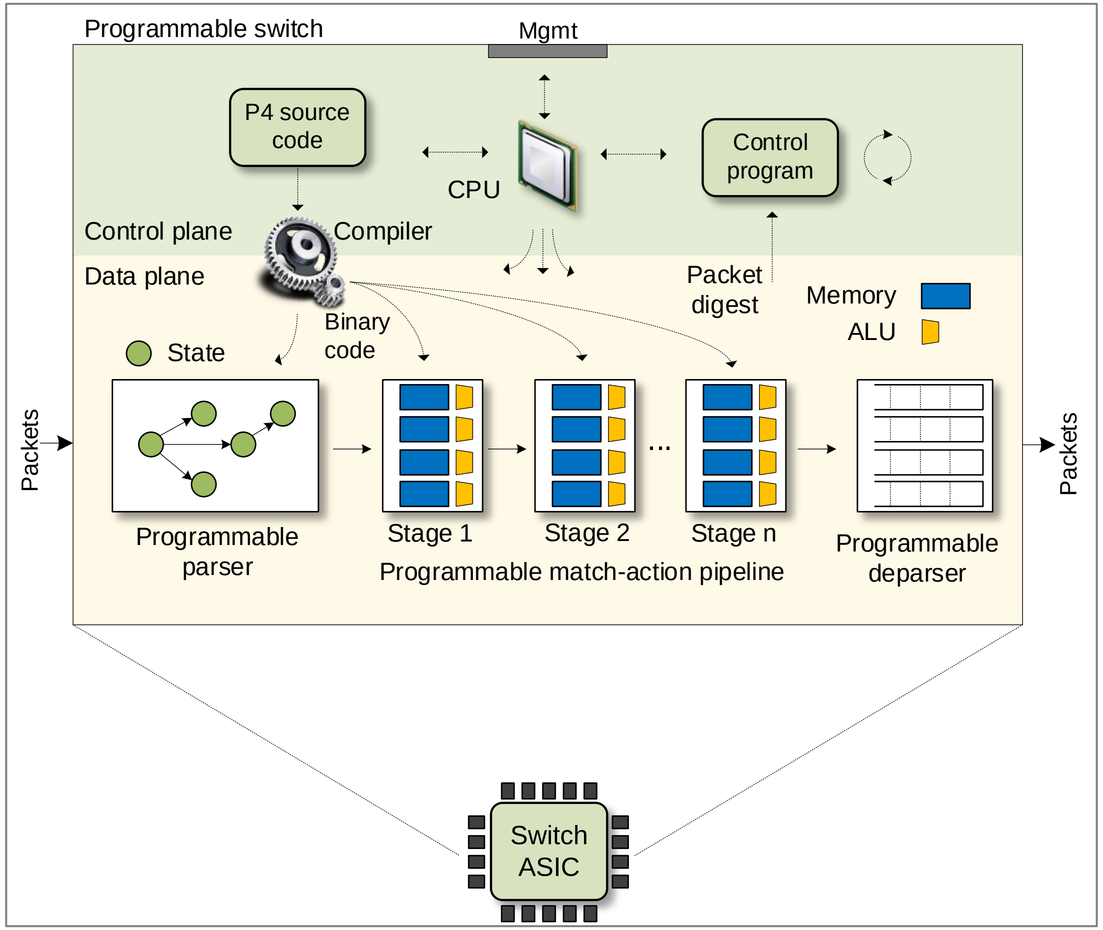
Exercise 2: Compiling and Testing a P4 Program

Exercise 3: Parsing UDP and RTP

Exercise 4: Building a Simplified NAT

Exercise 5: Configuring Tables at Runtime

Exercise 6: Building a Packet Reflector



P4 Applications and Custom Processing Lab Series

Lab experiments

Lab 1: Introduction to Mininet

Lab 2: Introduction to P4 and BMv2

Lab 3: P4 Program Building Blocks

Lab 4: Defining and processing custom headers

Lab 5: Monitoring the Switch's Queue using Standard Metadata

Lab 6: Collecting Queueing Statistics using a Header Stack

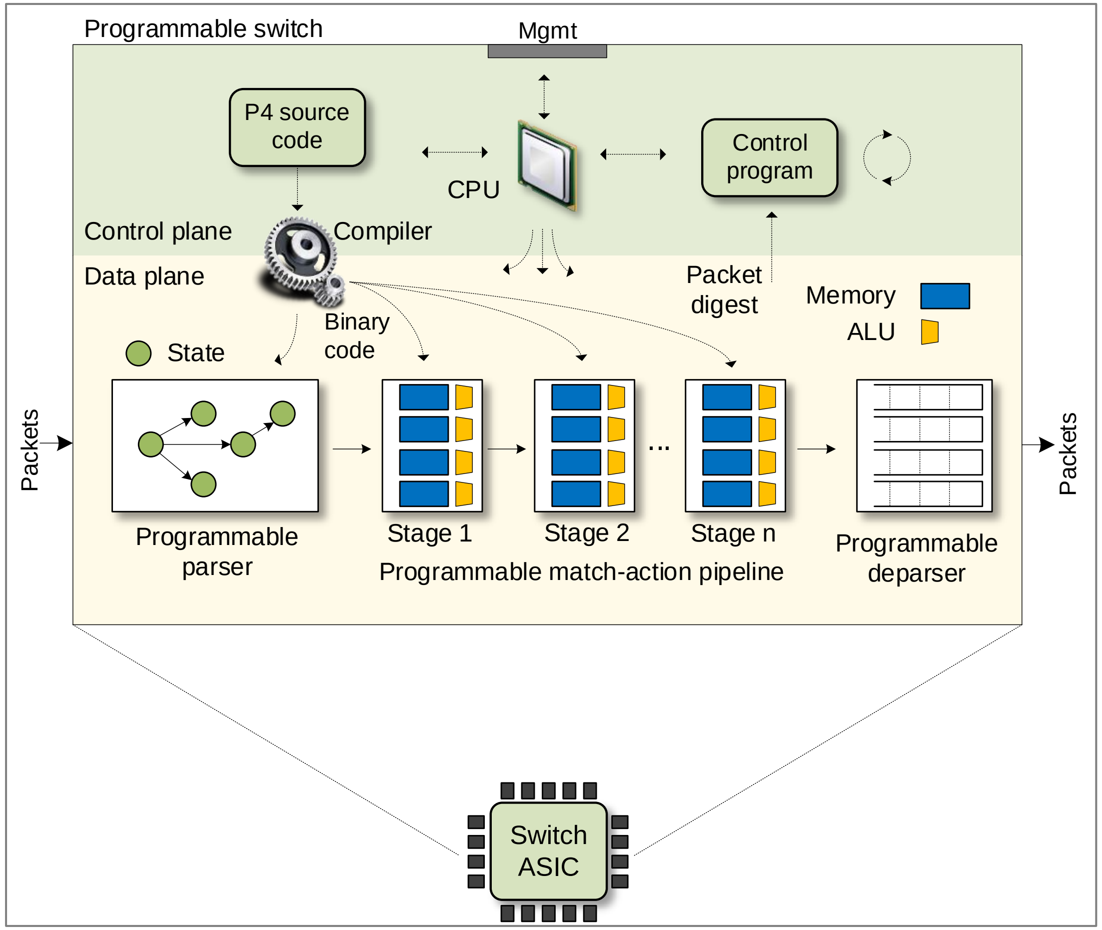
Lab 7: Measuring Flow Statistics using Direct and Indirect Counters

Lab 8: Rerouting Traffic using Meters

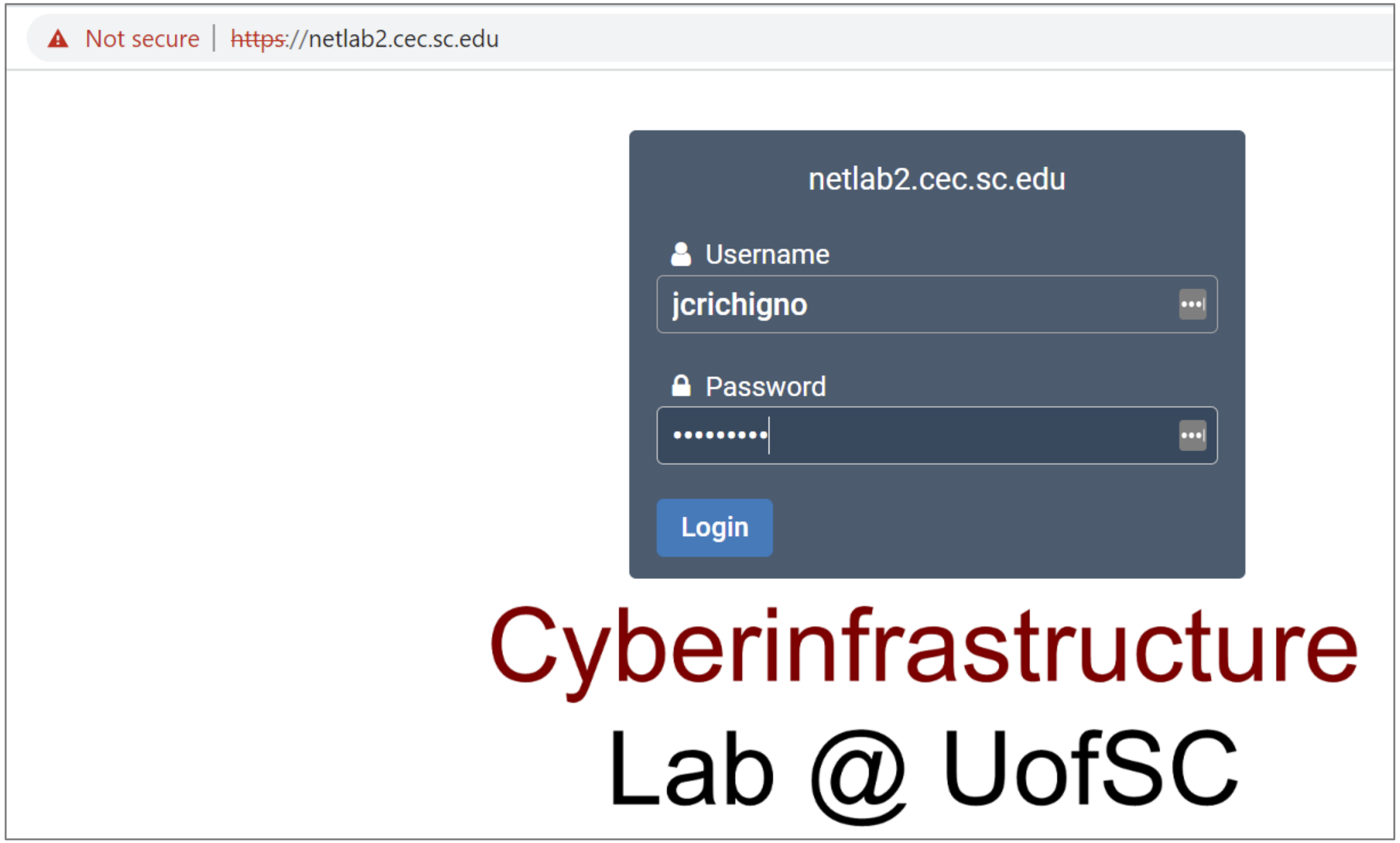
Lab 9: Storing Arbitrary Data using Registers

Lab 10: Calculating Packets Interarrival Time w/ Hashes and Registers

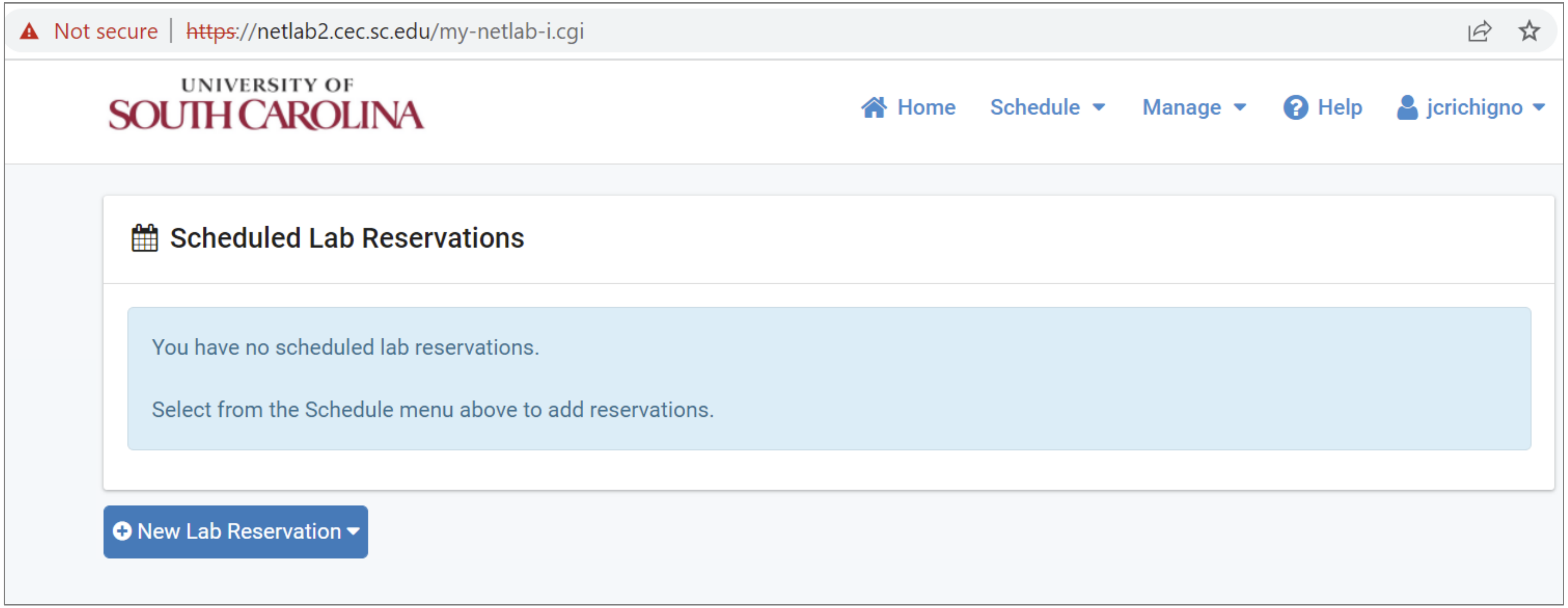
Lab 11: Generating Notification Messages from the Data Plane



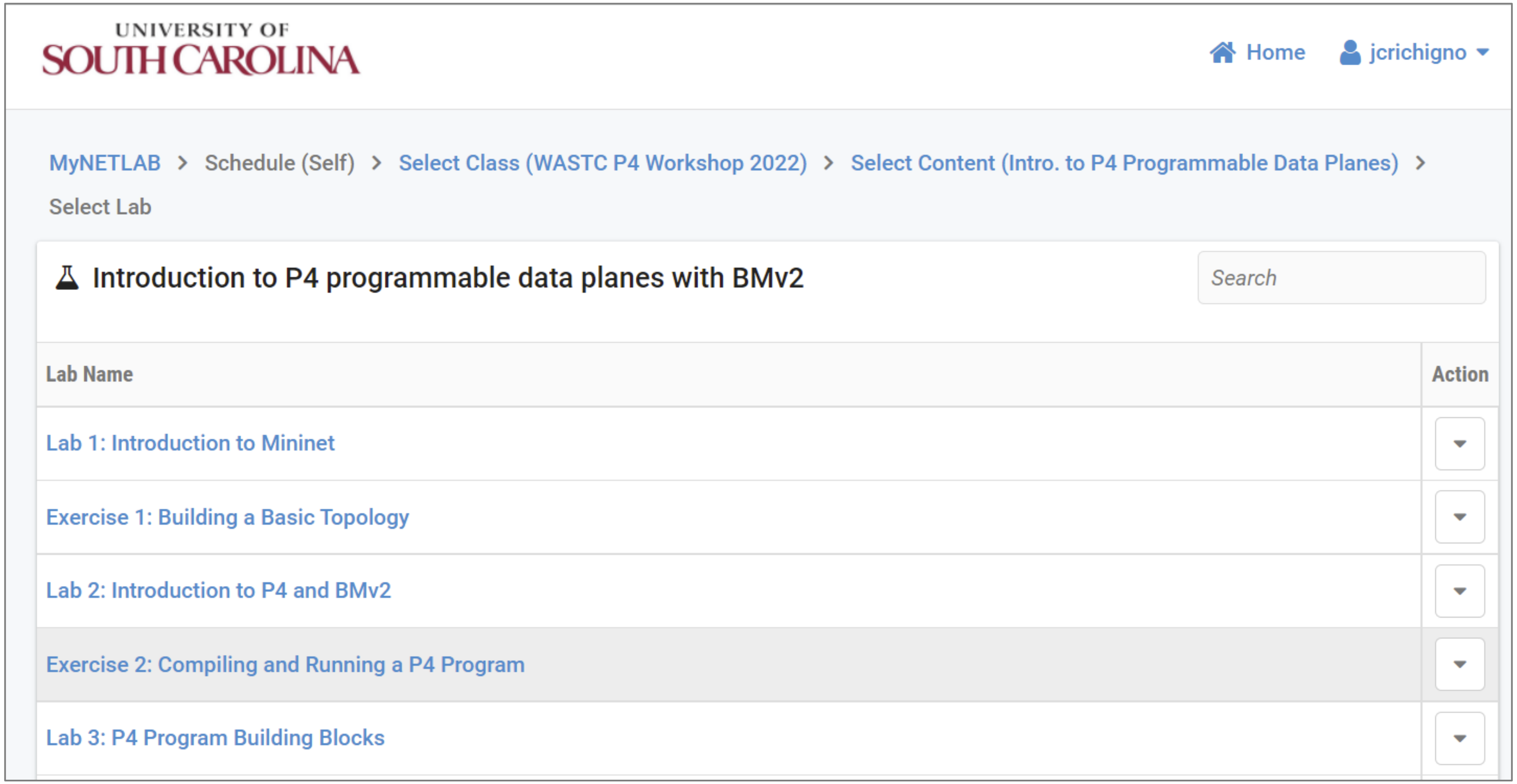
Using the Cloud System



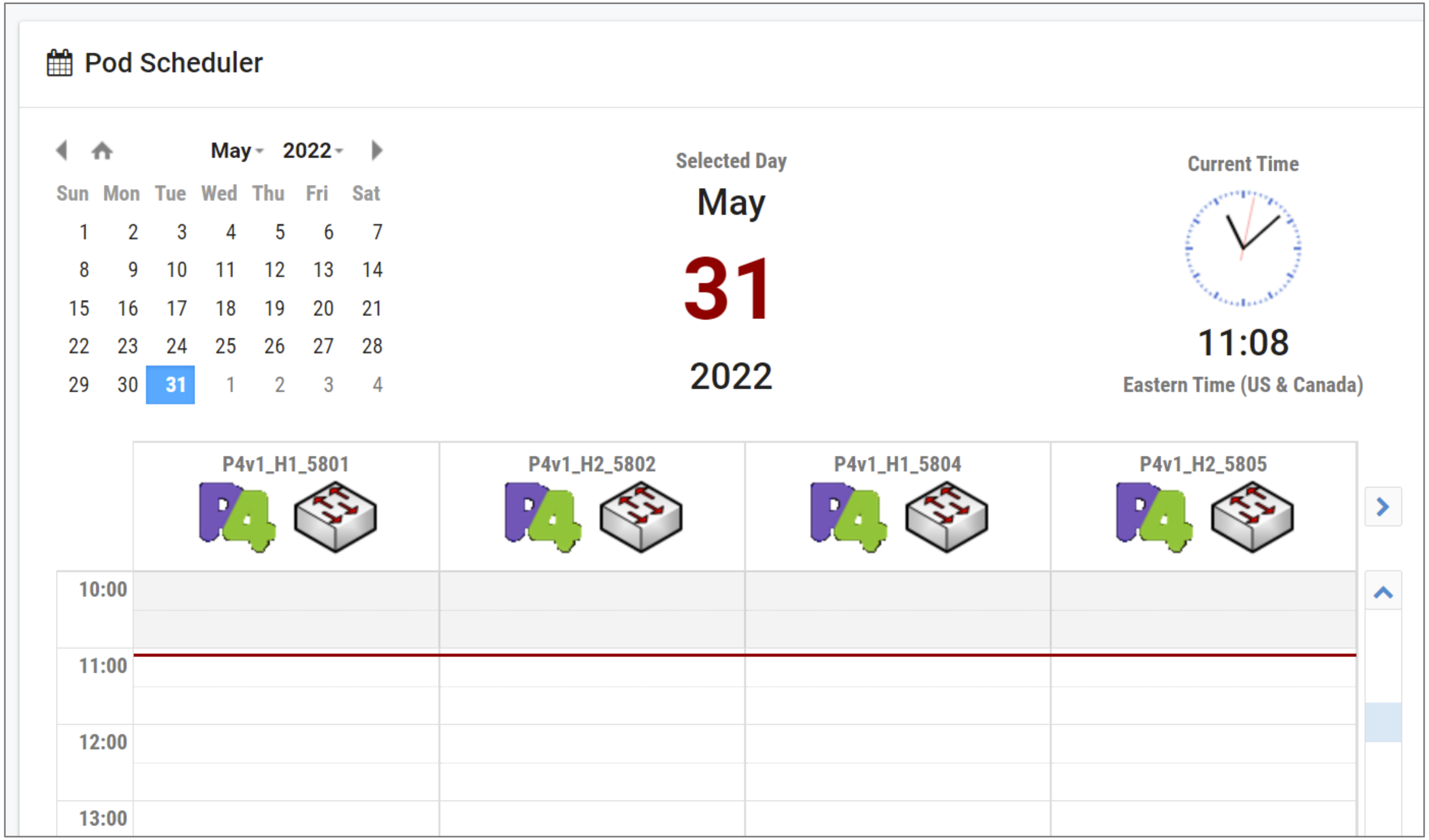
Using the Cloud System



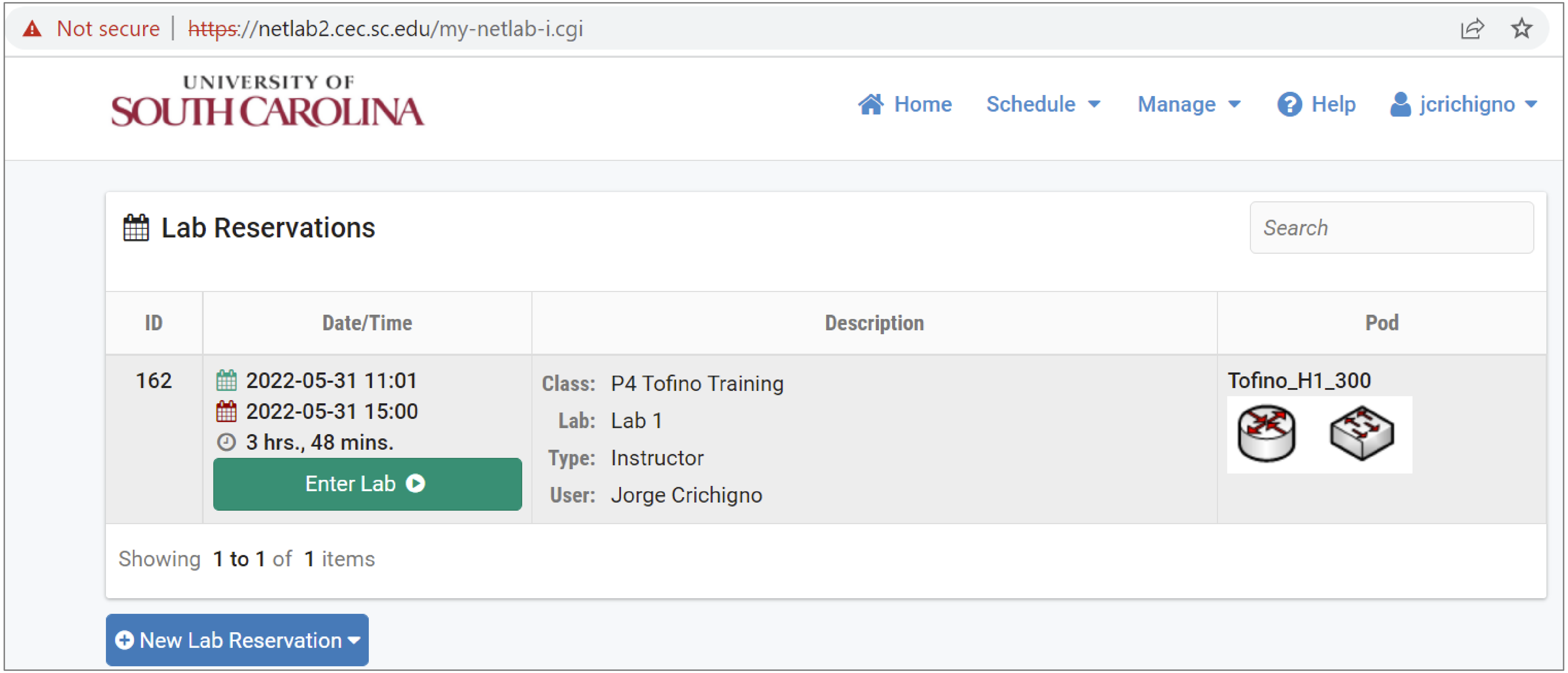
Using the Cloud System



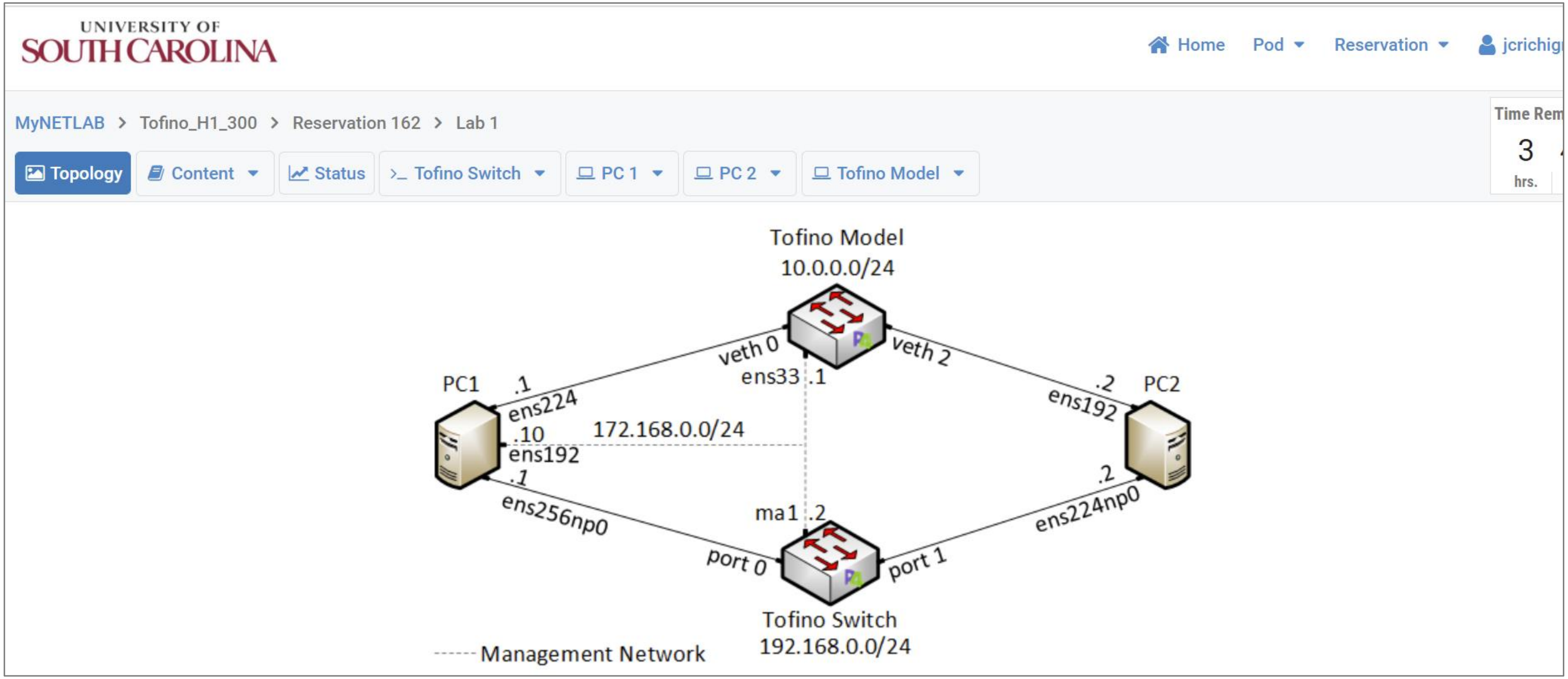
Using the Cloud System



Using the Cloud System



Using the Cloud System



Using the Cloud System

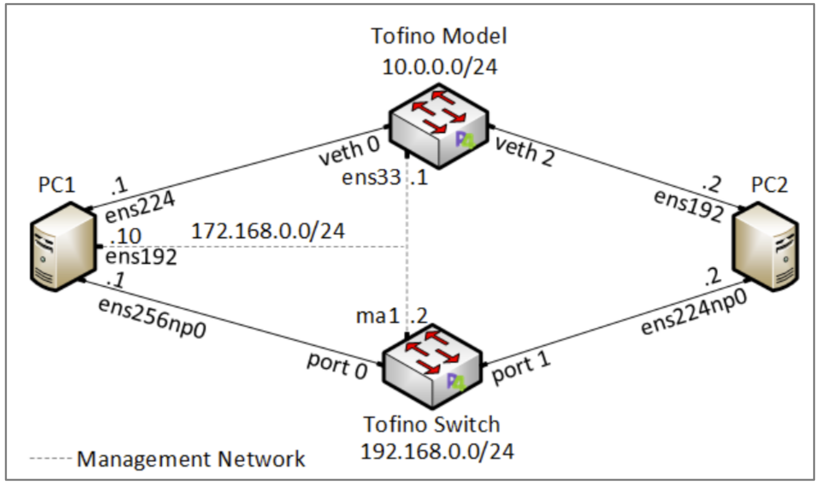
Using the Cloud System

- Readily available platform
- Topology complexity
  - ➢ 6.4 Tbps programmable switch
  - ➢ Tofino programmable chip (Intel)
  - ➢ Tofino model for debugging (trace execution in the data plane)
  - ➢ Servers to send/receive data to/from the switch/other servers
  - ➢ Multi-mode fiber
  - ➢ QSFP28+ transceivers
  - ➢ Open Network Linux (ONL) (control plane)
  - ➢ Software Development Environment (SDE) from Intel (control plane)
  - ➢ Compiler
  - ➢ Sample P4 codes for each lab (data plane)
  - ➢ Laboratory experiments with step-by-step directions (thousands of development hours)

⋮

- Logistics
  - ➢ NDA with Intel, lawyers' agreement
  - ➢ Procurement process
  - ➢ Physical hardware, rack space, data center, etc.
  - ➢ Software tools, SDE, operating system, etc.

⋮



Cloud Features

Feature	Comments
Allocation of resources	Pod granularity
Custom pods	Easy to create custom pods
Cost	Cost-effective when used extensively
Presentation layer for pedagogy	Topology is graphically presented to the learner using a regular browser
Time sharing	The owner controls who can access resources; easy to implement time-sharing policies
IP addresses	Pods (and learners) can have the same topology and IP addresses (overlapping addresses w/o conflict)
Functional realism	Virtual labs have the same functionality as real IT hardware in a real deployment, and execute the same code
Traffic realism	Devices generate/receive real, interactive network traffic to/from the Internet, or to/from other devices within the lab environment

Concluding Remarks

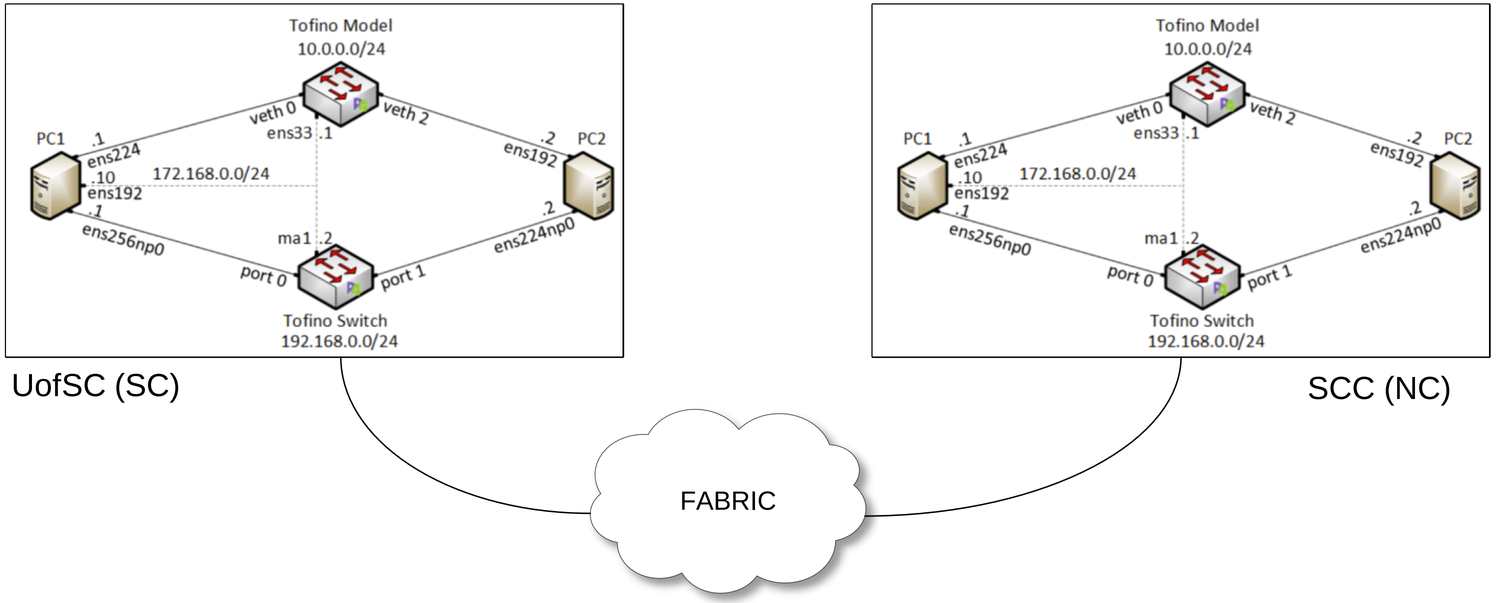
- The Academic Cloud has served over 100,000 learners (dozens of virtual libraries: Linux, virtualization, cybersecurity, etc.)
- Academic institutions (colleges, universities, high-schools), training centers
- Self-pace learners
- Usage example from one institution supporting one academic program (~300 students, January 1, 2020 – December 30, 2020)

Community Usage

Reservations Made	Labs Attended	Hours Reserved	Hours Attended
6909	6534	42446.59	25158.03
6909	6534	42446.59	25158.03
6909	6534	42446.59	25158.03

Concluding Remarks

- The system has shown to be scalable
  - It has served over 100,000 learners in 2020
- Due to the positive feedback, the system is expanding with more virtual labs
- The team is exploring the viability of connecting the Academic Cloud to FABRIC
- URL: http://ce.sc.edu/cyberinfra/cybertraining.html



Acknowledgement

- This work is supported by the National Science Foundation, award 2118311

UNIVERSITY OF
SOUTH CAROLINA

Academic Cloud vs Public Clouds

Feature	Academic Cloud	Public Cloud
Allocation of resources	Granular allocation of physical resources (CPUs, NICs, etc.)	Not granular (access to the physical resources requires additional fees)
Custom pods	Easy to create custom pods	Difficult; hard to design complex topologies
Cost	Cost-effective when used extensively	Cost-effective for individual / small VMs; costly for large VMs over time
Presentation layer for pedagogy	Very flexible. Topology is graphically presented to the learner using a regular browser	Not flexible; limited to providers' interface, e.g., command-line interface
Time-sharing resource feature	The owner controls who can access resources. Easy to implement time-sharing policies	Cloud provider controls who can access resources (typically, a fee is required per user)
Integration of physical devices	Easy; physical hardware can be integrated into pods	Difficult; no subscription plan permits integrating customized physical devices
Flexible use of IP addresses and subnets	Each pod runs in a sandbox. Pods (and learners) have the same topology and IP addresses (overlapping addresses without conflict)	IP addresses are typically unique. The vLabs manuals and companion material are not identical, requiring per-learner adjustment
Target	Specially built for pedagogy (education, research, and training)	General, used by a large variety of users
Typical users	From entry-level learners to PhD researchers	More experienced professionals, educators, students