



CloudLab





The Need Addressed by CloudLab

- Clouds are changing the way we look at a lot of problems
 - Giving us new ideas of what's possible
 - Impact goes far beyond computer science
- ... and have broader impacts with much more potential
 - Transformational for IT-based businesses – enables rapid startup
- ... but there's still a lot we don't know, from perspective of
 - Researchers (those who will transform the cloud)
 - Users (those who will use the cloud to transform their own fields)
- To investigate these questions, we need:
 - Flexible, scalable **scientific infrastructure**
 - That enables exploration of **fundamental** science in the cloud
 - Built **by** and **for** the research community



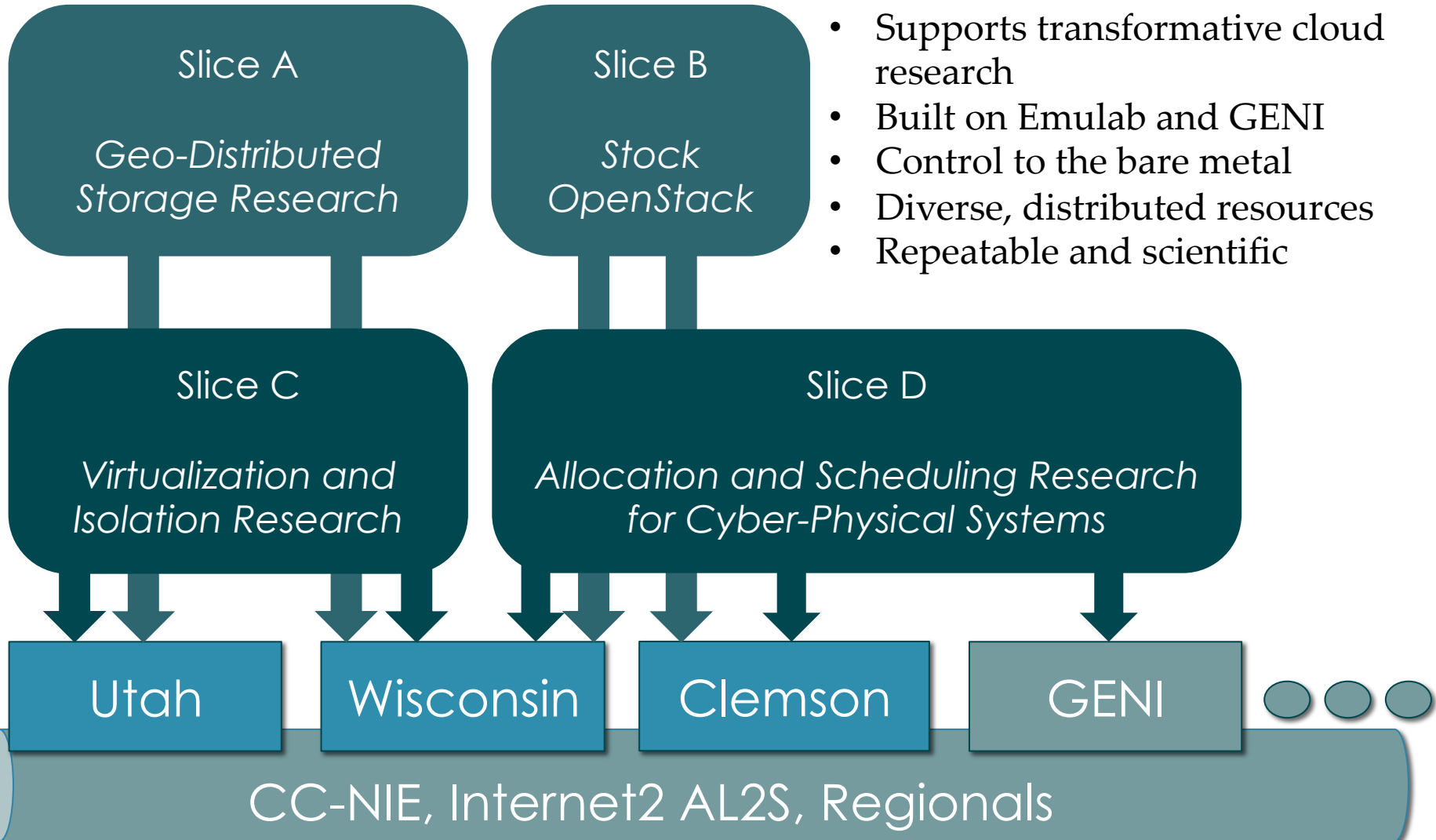
The CloudLab Vision

- A “meta-cloud” for building clouds
- Build your own cloud on our hardware resources
- Agnostic to specific cloud software
 - Run existing cloud software stacks (like OpenStack, Hadoop, etc.)
 - ... or new ones built from the ground up
- Control and visibility all the way to the bare metal
- “Sliceable” for multiple, isolated experiments at once

With CloudLab, it will be as easy to get a cloud tomorrow as it is to get a VM today



What Is CloudLab?





CloudLab's Hardware

One facility, one account, three locations

- About 5,000 cores each (15,000 total)
- 8-16 cores per node
- Baseline: 4GB RAM / node
- Latest virtualization hardware
- TOR / Core switching design
- 10 Gb to nodes, SDN
- 100 Gb to Internet2 AL2S
- *Partnerships with multiple vendors*

Wisconsin

- **Storage and net.**
- Per node:
 - 128 GB RAM
 - 2x1TB Disk
 - 400 GB SSD
- Clos fat-tree
- *Cisco*

Clemson

- **High-memory**
- 256 GB RAM / core
- 16 cores / node
- Bulk block store
- Net. up to 56Gb
- High capacity
- *Dell*

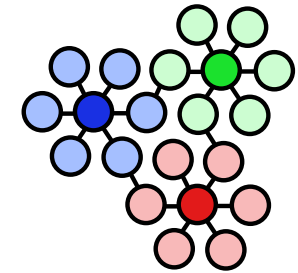
Utah

- **Power-efficient**
- ARM64 / x86
- Power monitors
- Flash on ARM64s
- Disk on x86
- Very dense
- *HP*



Technology Foundations

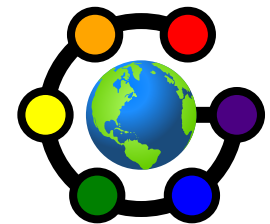
- Built on Emulab and GENI (“ProtoGENI”)
- In active development at Utah since 1999
- Several thousand users (incl. GENI users)
- Provisions, then gets out of the way
 - “Run-time” services are optional
- Controllable through a web interface and GENI APIs
- *Scientific instrument for repeatable research*
 - Physical isolation for most resources
 - *Profiles* capture everything needed for experiments
 - Software, data, and hardware details
 - Can be shared and published (eg. in papers)



emulab



geni
Exploring Networks
of the Future



protogeni



Who can use CloudLab?

- US academics and educators
 - Researchers in cloud architecture and novel cloud applications
 - Teaching classes, other training activities
- No charge: free for research and educational use
- International federations expected
- Apply on the website at www.cloudlab.us



Early Interest in CloudLab





Cloud Architecture Research

- Exploring **emerging and extreme** cloud architectures
- Evaluating **design choices** that exercise hardware and software capabilities
- Studying **geo-distributed** data centers for low-latency applications
- Developing different **isolation** models among tenants
- Quantifying **resilience** properties of architectures
- Developing new **diagnostic** frameworks
- Exploring cloud architectures for **cyber-physical systems**
- Enabling **realtime** and near-realtime compute services
- Enabling data-intensive computing (“**big data**”) at high performance in the cloud



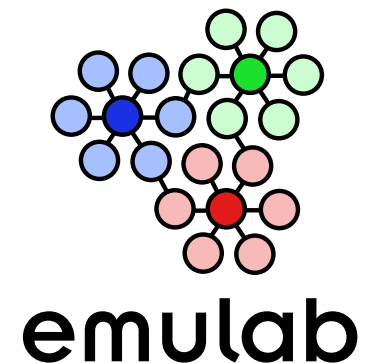
Application Research Questions

- Experiment with **resource allocation** and scheduling
- Develop enhancements to **big data frameworks**
- Intra- and inter-datacenter **traffic engineering** and routing
- New tenant-facing **abstractions**
- New **mechanisms** in support of cloud-based services
- Study adapting **next-generation stacks** to clouds
- New troubleshooting and **anomaly detection** frameworks
- Explore different degrees of **security** and isolation
- **Composing** services from heterogeneous clouds
- **Application-driven** cloud architectures



Federated with GENI

- *CloudLab can be used with a GENI account, and vice-versa*
- GENI Racks: ~ 50 small clusters around the country
- Programmable wide-area network
 - Openflow at dozens of sites
 - Connected in one layer 2 domain
- Large clusters (100s of nodes) at several sites
- Wireless and mobile
 - WiMax at 8 institutions
 - LTE / EPC testbed (“PhantomNet”) at Utah
- International partners
 - Europe (FIRE), Brazil, Japan





Community Outreach

- Applications in areas of national priority
 - Medicine, emergency response, smart grids, etc.
 - Through 
- “Opt in” to compute jobs from domain scientists



- Summer camps
 - Through Clemson data-intensive computing program
- Under-represented groups





Availability and Schedule

- Availability:
 - **Now (fall 2014): Technology preview available!**
 - Late 2014: Open to early adopters
 - *Early 2015: Publicly available*
- Hardware being deployed in stages:
 - Fall 2014: Dell / Clemson cluster
 - Winter 2014: Wisconsin / Cisco cluster
 - Spring 2015: Utah / HP cluster
- Hardware refreshes in 2015 and 2016



The CloudLab Team



Robert Ricci (PI)
Eric Eide
Steve Corbató
Kobus Van der Merwe



Aditya Akella (co-PI)
Remzi Arpaci-Dusseau
Miron Livny



KC Wang (co-PI)
Jim Bottum
Jim Pepin
Amy Apon



Chip Elliott (co-PI)
Larry Landweber



Mike Zink (co-PI)
David Irwin



Glenn Ricart (co-PI)





Learn more, sign up:

www.CloudLab.us



This material is based upon work supported by the National Science Foundation under Grant No. 1419199. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.