Report on the NSF Workshop on Next Generation Cloud Computing Research Infrastructure

November 11-12, 2019
Princeton University
Princeton, NJ

This activity was supported by Award CNS-1934704 and organized by Jack Brassil (Princeton University) and Mosharaf Chowdhury (University of Michigan).

Contents

Executive Summary

The CISE research community periodically assesses the specialized research infrastructure it needs to advance science. A 2-day workshop in November, 2019, in Princeton, NJ, convened 45 members of the community to engage in a dialogue about future directions for CISE Cloud Research Infrastructures (CRI).

The workshop had the following goals:

- assess the current state of cloud research infrastructure, and its role in the context of evolving academic use of commercial cloud offerings;

- provide a forum for infrastructure developers and experimenters to meet and share an objective and constructive dialog about future directions for cloud infrastructure;

- exchange ideas with industry participants -- primarily from commercial cloud technology providers and public cloud operators -- about the complementary role of public infrastructure to advance science, and their anticipated technology requirements; and

- assess the current and future impact that cloud research infrastructures can and will have on commercial cloud offerings.

The overarching workshop finding was that the CISE research community is well served by current CRI offerings. Additional key findings included:

- ✓ CRI and commercial cloud systems offer important and complementary services and capabilities;

- ✓ Participation in CRI development and experimentation by researchers from underserved institutions and underrepresented groups can be increased;

- ✓ Coordination with related international cloud research infrastructures and testbeds can be strengthened;

- ✓ The quantity of resources in the existing CRIs are not sufficient to meet peak community demand; and

- ✓ The diversity and capability of hardware resources in commercial cloud platforms are evolving, and should evolve in CRIs as well.

To ensure the broadest possible community input, an online survey of the CISE research community was undertaken in advance of the workshop to further understand current use and expectations of cloud computing infrastructures. The importance of CRI was reflected in the fact that 212 members of the community responded to the survey. Among the key survey findings were:

- ✓ Overall, the experimenter community reports being very satisfied with the existing CRI systems, with nearly 85% of responses rating satisfaction as 4/5 or 5/5 (5 is best);
- ✓ CRI supports the spectrum of computer and networking systems research topics;
- ✓ Active CRI experimenters invest considerable time in CRI experimentation;
- ✓ Commercial cloud services do not offer certain desired capabilities that can be found in some CRIs designed for science experimentation (e.g., transparency, experiment isolation and repeatability); and
- ✓ Most experimenters cite cost uncertainty as a significant barrier to experimentation on commercial cloud computing alternatives to CRI. An additional barrier for experimenters with legacy grants is their lack of support for paying cloud computing service charges.

Overall, the survey results indicate that the CISE research community finds current CRI systems meets critical community infrastructure needs, are well run, and provides a unique and important offering to the research community.

Introduction

Cloud Computing Research Infrastructure (CRI) provides an environment for computing and networking systems researchers to host their scientific experiments to advance knowledge and to educate the next generation of systems researchers. The academic research community has relied extensively on NSF CISE supported CRIs such as the NSF Future Cloud systems (CloudLab, Chameleon Cloud) to advance technologies underlying today's commercial cloud systems, including computer and network virtualization, Infrastructure-as-a-Service (IaaS) deployments, advanced programming models and software stacks, and software-defined networks.

Commercial compute clouds (e.g., AWS, GCP, Azure) have also evolved rapidly in recent years, and promise to play a growing role in accelerating scientific research. Public cloud platforms provide on-demand, elastic, and self-service access to hardware resources at scale, democratizing large-scale and/or big data computing. Public clouds provide easy access to modern generation computing hardware and advanced software stacks, allowing experimenters to quickly and cheaply investigate the latest technologies that are made available in the cloud. Commercial cloud platforms increasingly offer software environments to investigate important new data modalities such as streaming data and real-time analytics on such data. CISE researchers and educators can and do leverage these cloud platforms to accelerate and improve their research and teaching, complementing their use of CISE CRI systems and campus-based research computing facilities.

Given the evolution of research community experimental requirements, and the rapid evolution of commercial cloud offerings, the CISE research community faces the need to periodically assess the specialized research infrastructure it needs to advance science. A 2-day workshop was held on November 11-12, 2019, in Princeton, NJ, to convene the research community to engage in a dialogue about future directions for CISE cloud research infrastructures. The workshop brought together representatives from academia, industry, and government to articulate a research agenda for advancing cloud research infrastructure.

To ensure the broadest possible community input, the workshop was supplemented by an online survey of the CISE research community to further understand their current use and expectations of cloud computing infrastructures.

The workshop had the following goals:

- convene key cloud research infrastructure stakeholders to evaluate the current state of academic cloud research infrastructure, and its role in the context of evolving academic use of commercial cloud offerings;

- provide a forum for infrastructure developers and experimenters to meet and share an objective and constructive dialog about future directions for cloud infrastructure;
- exchange ideas with industry participants -- primarily from commercial cloud technology providers and public cloud operators -- about the complementary role of public infrastructure to advance science, and their anticipated technology requirements;
- allow researchers to exchange information about detailed system needs to address their future research experiments; and
- assess the current and future impact that cloud research infrastructures can and will have on commercial cloud offerings.

Background

Large-scale, data-intensive systems experimentation is increasingly important both for accelerating science and to understand the sorts of big data problems industry is currently engaging. Developing the most effective cloud research infrastructure will help ensure that learnings from the computing and networking systems research community will transfer successfully to commercial cloud systems.

The emergence of public cloud infrastructure over the past 15 years has led to innovations in distributed programming frameworks and systems, data center networking architectures, and advances in edge computing, machine learning, and 'big data' systems and applications. Academic scholarship has also been notably changed -- cloud computing has accelerated scientific discovery in fields such as genomics, high energy physics, and neuroscience.

Many of the innovations supporting cloud computing originated in the CISE research community. CISE supported the early development of cloud computing through investments in both foundational research and shared Community Computing Research Infrastructure (CCRI), including both Major Research Infrastructures (MRI) and Mid-scale Networking and Computing Research Infrastructures. Popular shared community systems including CloudLab [cloudlab], Chameleon Cloud [chameleon], and JetStream [jetstream] have provided the computing and networking systems community platforms to explore research investigations that would have been impractical using commercial systems.

As cloud computing research has evolved, NSF has continued to respond in various constructive ways to support cloud exploration by the systems research community. CISE has concurrently introduced several initiatives to ensure the acceleration of science through application of public cloud computing [ecas, bigdata, cloudaccess].

Commercial cloud computing has evolved rapidly in recent years. First, the growth and scale of commercial cloud enterprises has accelerated; the top providers are estimated to grow CAPEX investment at roughly 30% annually [google]. Second, cloud providers have sought product differentiation by expanding from their traditional commodity IaaS business to novel software offerings providing services ranging from analytics stacks to media service environments. Third, growing numbers of boutique cloud services have emerged, including those providing unconventional products such as the bare metal offerings of BigStep [bigstep] and Rackspace OnMetal [onmetal]. Fourth, new computation hardware including TPUs [tpu] and GPUs [ec2-gpu] have broadened boutique cloud appeal to new science disciplines. Fifth, new programming environments and tools including an emphasis on lightweight containers and Functions-as-a-Service (FaaS) [lambda] have been introduced.

As commercial cloud offerings have evolved and grown increasingly sophisticated, academic researchers have also moved forward in exploring multiple new research avenues in human-scale and human-centric computing, 5th generation cellular communications, and the Internet of Things. Many of these domains are believed to be enabled by future centralized and edge cloud computing architectures, supporting time- and safety-critical cyber-physical applications for medical devices, electrical power grids, and autonomous transportation systems.

The rapid evolution in cloud computing technologies and emerging research applications suggested a fresh look and updated community discussion of the requirements for future research infrastructures will help ensure continued community contributions to future cloud technologies. The Princeton workshop expanded a series of earlier workshops that have contributed to advancing a next-generation cloud computing research infrastructure aligned with a community vision of a future cloud systems research agenda.

Recent Workshops and Community Input

As community interest in cloud computing research has grown, a number of workshops and meetings have been held in recent years to advance the community's understanding of evolving cloud technologies. Many of these workshops were focused on specific aspects of cloud computing, and associated research infrastructure. The findings of these workshops informed our organization of Princeton Workshop topics. These community gatherings have included:

NSF Workshop on Grand Challenges in Edge Computing

Held in Washington, DC, in October, 2016, this workshop focused on edge computing and identified the following goals:

1) foster the edge computing community to increase interaction between academia and industry;
2) set the vision and identify grand challenges and open problems; and
3) identify collaboration mechanisms among academia, industry and government.

The workshop report [edgecloud] touched briefly on testbeds in asking "How can application developers develop applications that may eventually be deployed across multiple domains of different ownerships and levels of trust?" To address this question, a workshop finding stated that

> *"For researchers to fully explore edge computing, we need testbeds at scale."*

NSF Cloud 3.0 Workshop

The *Cloud 3.0* workshop [cloud3] was held at Stanford University in May, 2018. This was a focused workshop with the goal of bringing together leading researchers and practitioners in computer and networked systems to examine how emerging developments in cloud architecture and services could influence the community's research agenda. Among the motivations for the workshop was the relatively recent emergence of promising commercial cloud offerings in *serverless* computing and functions-as-a-service (e.g., AWS Lambda). This workshop report also addressed future testbeds:

> *"In building a testbed that supports FaaS research, it is useful to understand experiments where the scale is really important," where experiments using 1,000 cores was considered. It was also found that "Instrumentation and monitoring are crucial",*

and

> *"Today, on CloudLab and Chameleon, experimenters often want to be able to download open source version of something that is popular ... and then do research ... It would be valuable to enable something similar for Cloud 3.0."*

Other ideas related to testbed use in the workshop report included new datacenter architectures and networking interconnects, disaggregated network architectures, experiment isolation, and programmable fabrics and control planes.

NSF Workshop on Enabling CISE Research and Education in the Cloud

In Washington, DC, in January, 2018, a workshop [academiccloud] was convened

> *"... to explore the idea of an academic cloud, a [public] cloud that provides the set of services and capabilities that serve the unique needs, workloads, and users of the CISE community."*

This workshop explored the gap between the offerings of commercial cloud providers and the computational needs of the academic research community. While not addressing research infrastructure directly, the workshop encouraged broader uses of commercial clouds, where appropriate, as a complement to use of CISE cloud CRI.

Mid-Scale Experimental Research Infrastructure Forum (MERIF)

MERIF is a recently initiated community coordination group tasked with ensuring the federation and interoperability of computing and networking research testbeds, including cloud research infrastructure. Beginning in May, 2019, MERIF is convening periodic research community meetings to discuss the current and future interoperation of systems. It is anticipated that the forum will work to crystallize learnings about community needs for future research infrastructures, including those from the Princeton workshop.

Future Directions for NSF Advanced Computing Infrastructure

Developing cloud research infrastructure is not performed in isolation; researchers flexibly use commercial cloud systems, campus research computing facilities, and laboratory based systems to perform their work. As a result, recommendations for cloud research infrastructure are informed by broader community guidance, including the National Academies report on Future Directions for NSF Advanced Computing Infrastructure [nationalacademies].
A particular recommendation of that broader activity was:

> *Recommendation 2.3: NSF should (a) eliminate barriers to cost-effective academic use of the commercial cloud and (b) carefully evaluate the full cost and other attributes (e.g., productivity and match to science workflows) of all services and infrastructure models to determine whether such services can supply resources that meet the science needs of segments of the community in the most effective ways.*

Many other focused meetings and/or workshops on Cloud Computing have been held in recent years, often narrowly focused on either cloud education or outreach [cloudlab4all, cloudinbox], application to a particular scientific discipline [cluster], or a specific infrastructure [cord, osg, opencloud]. None of these workshops have directly examined the future of mid-scale cloud initiatives, however in 2017 an independent research community group assessed the progress [krieger] of NSF Future Cloud [nsfcloud] infrastructures for the research community's own benefit.

Given the absence of a recent, broad assessment of future directions in cloud computing research infrastructures, and given the rapid rate of change of commercial cloud computing capabilities, we believe that the moment has come to convene the research community to explore next generation infrastructures that will best serve the computing and networking systems research community.

Key Workshop Findings

The workshop concluded with a plenary session to discuss major findings following 1.5 days of participant presentations, breakout sessions, and engaging question and answer sessions. This session provided an opportunity for participants to revisit and capture key learnings from previous discussions and propose ideas to further strengthen and broaden community use of CRIs. The following findings were discussed:

***Finding 1**: Research community is well served by current CRI offerings.*

The community reports both relying on and being highly satisfied with current CRI systems. This is supported by many measures. The experimenter community is large, with approximately 5000 current and previous users registered on the combined CloudLab and Chameleon user mailing lists. A wide range of domain science and computer systems research topics (e.g., machine learning, databases, applications, networking, systems, etc.) are benefitting from active experimentation on CRIs. Scholarly publications at prestigious academic conferences report experimental results created on CRI. While research is the predominant use of CRI systems, educational uses (e.g., class projects, homework assignments) are material and growing.

***Finding 2**: CRI and commercial cloud systems offer important and complementary services and capabilities.*

Experimenters recognize and acknowledge the distinct benefits and use of both commercial cloud and CRI systems. In particular, they value CRIs for providing a high degree of experiment isolation, support for experiment replicability, and critical services and resources such as bare metal machine access. CRI users also see resource scalability, availability, and software services as important properties of commercial cloud systems. But concerns about export control, cost uncertainty, payment with legacy grants, vendor lock-in, and “openness” are all cited as barriers to commercial cloud use. The experimenter view of both systems can be characterized as “distinct tools in the researcher’s toolkit”. Researchers select one or the other tool as appropriate for the research task. Few researchers articulate a vision of a deeper integration of these systems to support future experimentation.

***Finding 3**: Participation in CRI development and experimentation by researchers from underserved institutions and underrepresented groups can be increased.*

While the research community recognizes the potential for CRI to democratize research, they acknowledge that underserved institutions and underrepresented groups are not well represented among active CRI users. The composition of workshop attendees reflected this fact. Several participants pointed to previous successful initiatives including *CloudLab for Everyone* [cloudlab4all] as models for future outreach efforts. Attendees suggested that future approaches to be explored to broaden participation might involve a sustained commitment to funding outreach activities, and increased opportunities (e.g., internships, fellowships) for exposing students or faculty researchers from MSI/HBCUs on CRI research or development project teams.

***Finding 4**: Coordination with related international cloud research infrastructures and testbeds can be strengthened.*

Workshop attendees lauded the success of previous international efforts to co-develop, federate, or exchange information or artifacts from large research infrastructures. Examples of such activities include PlanetLab Europe, GENI/FIRE workshops, and the Global Environment for Future Infrastructure (GEFI) meetings. The workshop attendee composition reflected this lack of international representation; only a single participant was a non-US person.

Finding 5: The quantity of resources in the existing CRIs are not sufficient to meet peak community demand.

Given the popularity of the existing CRIs, the physical resources they have are not always sufficient to meet user demand, especially approaching major publication deadlines and other busy times during the academic year. Greater capacity in the CRIs is likely to be well-utilized and to improve the quantity and quality of research and education use of the platforms.

Finding 6: The diversity and capability of hardware resources in commercial cloud platforms are evolving, and should evolve in CRIs as well.

New technologies such as faster storage devices, new CPU generations, and programmable networks are rapidly introduced into commercial clouds platforms. To support the highest quality research and education uses, CRIs should also strive to introduce new technologies, while retaining their complementary capabilities and services. CRIs should also seek ways to offer resources of unique value to the research community that are unlikely to be supported by commercial providers. Potential ways to pursue this include supporting "bring your own device" schemes, and developing mechanisms for CRI users to indicate new resources to support their experiments.

Workshop Agenda

Day 1: Mon., November 11, 2019 – Nassau Inn

8:30-9:00am *Breakfast*
9:00-9:30am *Introduction*

· Introduction & Logistics: Jack Brassil, Mosharaf Chowdhury (10 min)
· Princeton University Welcome: Jen Rexford (5 min)
· Workshop goals: Deep Medhi, NSF (10 min)
· Participant Introductions (5 min)

9:30-10:30am ***Plenary I*** – *Cloud RI operators*

- *Chameleon: From Testbed to Ecosystem* – Kate Keahey (10 min)
- *Why We Need Testbeds for Cloud Computing* – Rob Ricci (10 min)
- *National and Local Impacts of Programmable CyberInfrastructure using Jetstream* -- David Hancock (10 min)

· Lightning Talks - KC, Manu, Timur, Peter/Orran, Dan, Mike (30 min)

- *Clouds and Internet Converged for the Future* – K.C. Wang (3 min)
- *Wireless Ecosystem meets the Cloud* – Manu Gosain (3 min)
- *EdgeNet: A Lightweight, Scalable Edge Cloud* – Timur Friedman (3 min)
- *The Mass Open Cloud (MOC)* – Peter Desnoyers / Orran Krieger (3 min)
- *Open Cloud Testbed*– Mike Zink (3 min)
- *TACC Insights* Dan Stanzione (3 min)

10:30-11:00am *Break*
11:00-12:15pm ***Breakout Sessions 1a-e***: Assessing where we are today: Cloud RI mechanisms, Application domains, Technology transfer, Cloud Integration, Cloud Training, Barriers to use

a) Cloud RI mechanisms – Federation, HW/SW, schedulers, diversity – what's already there, what works (and doesn't), what is needed, what isn't
b) Application domains – What research domains are supported, what is not? Edge, wireless? What is worth supporting?
c) Technology transfer – Identify successes and failures. Are we helping operators? Should that be a goal?
d) Cloud RI Education & Training – How are we and our students learning? Workshops? What works? What is the community health? Are we stimulating community engagement?
e) What are the barriers to academic use of Cloud RI? Ease of use? Should we invest more in documentation, professional software developers?

12:15am-1:15pm *Lunch*
1:15-2:00pm *Groups 1a-e report back*

2:00pm-3:00pm ***Plenary II***: How do Commercial Clouds fit in?

· *Operating Large Scale Virtualized WAN on Open Network Emulator* (Microsoft Azure) – Lifen Wu (15 min)

· Cloud Heavy Hitter User (Princeton Neuroscience) – Sebastian Seung, Will Silversmith (15 min)

· Lightning Round – Public Clouds: Tamar E, Rob. F., Jamie S. (12 min)

- *CloudBank* – Rob Fatland (3 min)
- *ECAS – Exploring Clouds for Acceleration of Science* – Jamie Sunderland (3 min)
- *Next Generation Cloud Infrastructure and Platform for Mission Critical Enterprise Applications* – Tamar Eliam (3 min)

3:00pm-3:30pm *Break*

3:30 – 4:00 ***Plenary III*** – Visions of Future Cloud RI

· Lightning Round: Jeff, Rick, Joe, Justin, Glenn (30 min)

- *Tiered Cloud* – Rick McGeer (3 min)
- *Should NSF Really Host Cloud Infrastructure?* – Jeff de La Beaujardière (3 min)
- *Desired Capabilities Of Next Generation Shared Experimental Cloud Computing Research Infrastructures* – *Joe Mambretti (3 min)*
- *Next Gen Cloud from the Science of Networked Computing* – Justin Shi (3 min)
- *Solid Clouds* – *Clouds Quick and Dependable Enough to be a Seamless Part of Reality* – Glenn Ricart (3 min)

4:00-5:15pm ***Breakout Sessions 2a-e***: *Future Cloud RI requirements*

a) Education & reaching underserved institutions – are we democratizing? Are special tools/practices needed?

b) Goals for future Cloud RI systems – What must be built to handle experiments in 5-10 years?

c+d) Does the research community just need bare metal as a service? HW Infrastructure –What compute, storage, and networking is needed?

e) What are the barriers to academic use of commercial cloud systems for Cloud RI users?

5:15-5:30pm ***Groups 2a-e report back***

6:00 – 8:30 pm – *Workshop Dinner*: Palmer House, 1 Bayard Ln. (use Nassau St. entrance)

Day 2: Tues., November 12, 2019 – Nassau Inn

8:30-9:00am *Breakfast*

9:00-9:15am ***Recap from Day 1***

9:15– 10:15am ***Plenary IV:*** Cloud RI Technology and Mechanisms (Analytics, Federation, SW/HW, etc)

· *Fabric Overview*: KC Wang/Jim Griffioen (10 min)

· Lightning Round: Brent, Mosharaf, DK, Chris, Jim, Chip, Paul, Amr., Mohammed, Stephen (50 min)

- *Choosing the Right Programmable NICs for Future Shared Cloud Infrastructures* – Brent Stephens (3 min)

- *Designing and Deploying High-Performance and Scalable MPI Library for Amazon-AWS and Microsoft-Azure Cloud* – DK Panda (3 min)
- *Experiments in integrating commercial cloud services with on premise research computing solutions* – Chris Hill (3 min)
- *Leveraging Federation to Enable New Types of Applications* – Jim Griffioen (3 min)
- *Federation – Chip Elliott (3 min)*
- *Computing across Clouds, Campuses, and Scientific Facilities* – Paul Ruth (3 min)
- *Data Management in an Untrusted Cloud – Amr El Abbadi (3 min)*
- *MergeTB – Stephen Schwab (3 min)*

10:15 – 10:45am *Break*

10:45-12:00pm ***Breakout Session 3***: Tools for Experimenters

a) All Things Data - management, virtualization, sharing, infrastructure data use
b) Federation - how to handle experiments spanning multi-cloud, multi-instrument
c) Experiment monitoring, debugging, understanding; Infrastructure monitoring & debugging
d) Future Cloud Research Infrastructure revisited

12:00-1:00 *Lunch*

1:00-1:30pm ***Breakout Groups 3a-d Report Back***

1:30 pm – 2:15 *Plenary V*: Big Science

· *Lightning Round – Science users: Kevin, Mark C, Ryan, Joe, Mohan, Anirban (30 min)*
- *Enabling Open and Reproducible Science using Cloud Research Infrastructure* -- Mohan Ramamurthy
- *Cloud Computing at NASA's Frontier Development Lab* – Mark Cheung
- *Cloud-Native Data Formats for Big Scientific Data* – Ryan Abernathey
- *Seven principles for effective scientific big data systems* – Joe Hamman
- *They did it and I reproduced it! Practical Scientific Reproducibility in the Cloud* – Kevin Tyle
- *Leveraging Distributed Cloud Testbeds for Domain Science Research and Experimentation* – Anirban Mandal

2**:15-3:30pm *Plenary V***: Group Discussion/Silent Group Google Doc Q&A

Workshop Google doc

3:30pm *Adjourn & disband*

3:45-5pm *Editorial Session*

Participants’ Report on Breakout Sessions

Breakout Session 1

Participants in this breakout session were charged with reflecting on and broadly assessing the state of existing cloud research infrastructures. The session was divided into 4 breakout groups focusing on different aspects and measures of health of existing CRI.

a. Cloud RI Mechanisms

Participants: Lawrence Landweber, Robert Ricci, Glenn Ricart, Paul Michael Ruth, Li-Fen Wu, Irene Kopaliani, Joe Mambretti

Participants in this breakout session were charged with assessing the state of the low-level hardware and software infrastructure mechanisms that form the foundation of existing cloud research infrastructures.

Federation

A discussion of federated infrastructure began with a rough consensus of a definition: the federation of facilities involves providing cross-facility access to the following resources: Storage, Network, Instrumentation/Measurement/Monitoring, and Data resources. The latter resource can include Databases, Repositories, System measurements, and Usage data (e.g., number of CRI users, number of experiments, etc).

The priority to associate with federation activities was debated; how important is it to federate infrastructure? Operators look to users to ultimately decide what resources they want to federate, which can include elements such as local instruments, fabrics, or other shared NSF-funded infrastructures.

Recommendation: Dedicated individuals or entities should to be responsible for infrastructure federation; cannot be accomplished on the margins

Finding: Federation at the authentication level (e.g., in support of single sign-on processes) mostly works well today. Other types of federation other layers are still significant struggles (authorization and allocations).

Research Challenge: How or can you federate measurement infrastructure? Are there common attributes that can be shared or exported? For a lot of experiments the measurement framework is core to the problem.

What is missing?

Experimenters continue to express interest in bring-your-own-technologies to enhance CRI. The list of possible technologies goes well beyond devices (i.e., BYOD), and includes hardware, synchronization, message passing, reliability, security, orchestration, etc.

Research challenge: To what extent can a bring-your-own model be realized on a shared infrastructure (while continuing to retain support for resource sharing and experiment isolation)?

For example, PAWR provides for a single user of a whole testbed for limited periods of time which allows a `bring your own' anything. But even on shared infrastructures, researchers should be able to bring their own communication failure models, varying latency, potential adversary attacks, etc. This concept is powerful because it leads to the concept of a user-extensible testbed. It may be possible to reduce the pressure to create brand-new testbeds to serve specific needs if the existing testbeds become more extendible. The Massachusetts Open Cloud extension of the CloudLab testbed might be viewed as a successful first opportunity to examine this opportunity.

Recommendation: Future NSF CRI solicitations can challenge the research community to explore approaches and mechanisms, and formal processes to support infrastructure extensions.

The participants also identified a list of other technologies and processes that are either missing or not fully developed in today's CRIs, including: technologies (SSDs, NICS, NVIDIA single-precision floating point, Machine learning accelerators: TPUs) business models (a PlanetLab like co-op model), super-facilities (federation with specific data repositories or scientific instruments), and data resources (e.g., traces).

b. Application Domains

Participants: Chip Elliott, Abhimanyu Gosai, Stephen Schwab, Ann Von Lehmen, Timur Friedman, Brent Stephens, Anriban Mandal, Justin Shi, Ryan Abernathy, Chris Hill

This discussion sought to identify what research domains are currently well supported by existing CRIs, and what are not. Examples of domains could include Computer Systems research such as edge computing or wireless communications, or domain sciences such as Geoscience. What application domains should be supported?

The group observed that domain scientists want or need a pure service that provides end-to-end cloud computing workflows. These domain science groups often struggle with the current model of interacting with commercial cloud systems, where the burden of cloud infrastructure falls on small research teams to build and sustain their infrastructure (as opposed to focusing on the science). This model is effective, but the resulting infrastructure is hard to sustain. One challenge with certain domain science applications is that multiple funding agencies might be supporting a project. A second challenge is the reality that considerable science data is distributed, and moving data (e.g., between different cloud providers) is expensive. These factors raise questions about the continuity of funding for a project's cloud services & storage.

Recommendation: Can funding agencies provide support for to sustained infrastructure operations for this class of science domain user?

APIs and Services

The group discussed an aspirational service abstraction to support domain science applications. An end-to-end API would allow the best-effort performance and reliability using available networks, processors and storage. The service would define the actual infrastructures that end-users see and rely on. Different end-to-end APIs would be developed for different abstraction layers. Validating such a service would be done by recruiting some lead users in the community to serve as early adopters.

The opportunity provided by such a service would be the possibility that workflows could be defined in a common way (e.g. Hadoop job such as CYVERSE @ U of Arizona) and deploy across commercial clouds (but also make them available to research cloud infrastructures). Ideally, the implementation team would think of sustainability from Day 1, and cross-pollination would benefit domain scientists. Capturing such public interfaces and workflows would also serve as interesting uses for computer and network scientists.

c. Technology Transfer

Participants: DK Panda, Orran Krieger, Jack Brassil, Dan Stanzione, Erica Lam, Tamar Eilam

This discussion sought to assess the current state of technology transfers from both advances in cloud infrastructure, and experiments performed on CRIs. The first challenge confronted was the scope of contributions that might be considered a tech transfer, and whether they are defined by hardware or software artifacts, methods or the recipient of those components. Examples discussed included open source software contributions to large software projects, as well as innovations such as application of FPGAs in the cloud.

Measuring the extent of a contribution is often difficult; while assessing the value of a contribution to an open-source project might be are hard to measure, other measures such as counting deployments in commercial systems, or licensing instances (or value) might be straightforward.

The quality of the transferred artifact remains a factor in evaluating transfer success; the code created in academic settings is rarely professional or near-professional quality. Industry engagement early in the transfer process is considered crucial; the Red Hat Collaboratory model of transfer was notable, where graduate students stop contributing at some point in the software development process. Early industry engagement is seen as a key success indicator; the group sensed an opportunity to build this engagement around testbeds. Early adopters seem to be a key as well. Industry engagement (early) is a requirement — how do we build them around testbeds.

While industry has provided some funding, government support is dominant. But possible approaches for how industry to contribute include access charges, or in-kind contributions of resources. In some cases, it is unclear how testbeds would help industry partners (e.g., IBM, Microsoft) directly – perhaps their cloud offerings can enable research better.

The group saw transfers include FPGAs, Ceph, Unikernel, Key Lime and Elastic Scalable Infrastructure as transfer models. The group also discussed the history of examples of value-adding ‘industry influencers’ rather than directly observable, tangible transfers. They noted this “soft impact” associated with the evolution from the publication of Mesos through the Borg release to Kubernetes today, as well as the role of Xen in helping move industry from proprietary KVM solutions to more open solutions.

The time required to execute a successful transfer was discussed. The MVAPICH High Performance open-source MPI Library was under development over a decade and a half before deployment in Azure, AWS, and Oracle Cloud. Was this an outlier or is long term engagement required? The group noted that the time and effort investment in tech transfers has “cost a fair number of publications.”

Recommendation: Industry can assist in identifying the value received from academic tech transfers.

One model is Microsoft research metrics of "product impact" that seeks to ensure that Azure and the internal research lab work closely (i.e. adoption of FPGAs).

A discussion on the inhibitors to using commercial cloud for academic researchers – in this case specifically considering research on algorithms – revealed the following inhibitors:

1.) cost;
2.) data transfer costs (to promote lock in) contrary to research ethos;
3.) cutting edge infrastructure access requires you to be "internal" — send student to work there is the only vector; and
4.) lack of transparency in information.

Participants discussed the differences inherent in studying or experimenting on an "operating cloud" versus a CRI or testbed. In some cases, industry researchers could benefit from access to an "open" "operating cloud" for access to data, service evaluation, etc. Interoperability between testbeds and commercial clouds could also support tech transfer. Uniform metric collection would be a positive. Reproducibility is an important requirement to build industry trust.

The Red Hat/Boston model was put forward as a good example of how to build these academic -industry partnerships. The ideas typically originate with researcher ideas and proposals to initiate join projects. Even then it is hard to get a commitment on the company side of a portion of a developer to work with the project to facilitate transfer. Starting work together from inception is more likely to gain traction than a project "thrown over the fence".

Challenges with the open source model were discussed — the existence of a github repo is not sufficient to build a community. How do we train/incentivize people to do build communities, a process academics often don't want to pursue? The importance of a community reputation to convince a community to follow you was also noted.

A proposed interesting approach to studying technology transfers was to ask, "What are failure examples?" Telemetry on MOC was cited as a possible example —there is no way to report telemetry in a standard way and the extant tools didn't work at scale. Researchers are frustrated that they cannot get usable data, and a series of industry proposals continues without a solution emerging.

d. Cloud RI Education & Training

Participants: Mosharaf Chowdhury, Regina Hain, Mike Zink, KC Wang, Rob Fatland, Amr El Abbadi, Mohan Ramamurthy

Participant discussed how and where CRI is working and not working to support the education mission. One participant shared that getting scale on CRI is challenging, esp. for longer period of time for class assignments for 40 graduate students. Commercial cloud came up as a possible solution, but it was argued that the mechanics of cloud use is hard with students.

The availability of GPUs (particularly in Chameleon) is a positive factor for education uses including machine learning applications. The group noted that work and new approaches are needed to educate/engage students in the deployment and operation of the systems

Commercial clouds are playing a growing role though support projects such as CloudBank. One education use today is support of data science classes like Berkeley Data 8.

The group noted the important role of JupyterHub in supporting various science education projects (e.g., PanGeo), and the role of various technologies for scaling cloud resources up and down (e.g., Zero to JupyterHub with Kubernetes, Google co-lab, etc).

The group also mentioned the wide range of computer systems education activities well supported by current CRI including data bases, fault tolerance, distributed systems, placement protocols, detect deadlocks because of real diverse network latency, edge computing, etc.

The participants also discussed opportunities to further promote educational uses of CRI

Recommendation: Education outreach in the form of events such as webinars and hackathons can help broaden student CRI use.

The sort of education and training needed for educating users (including students) about CRI and commercial cloud is varied. This can include basic know-how about cloud technologies, or about what problems are well suited to each platform. A common, standard ‘education environment’ might be of value. Use of multiple providers, multiple provider locations, distributed data, privacy issues, data movement legal issues, and funding could be components. The CloudBank model or 1) teaching how to use public cloud, and 2) helping to develop new research projects is only one approach. Improving understanding of different funding mechanisms including the ‘campus condo” model [future-aci] might also be a valuable topic for educating users.

e. Barriers to Academic Use of Cloud RI

Participants: Joe Hamman, Eric Eide, Kate Keahey, Kevin Tyle, Rick McGeer, Jeff de La Beaujardière

The group sought to identify and remedy existing barriers to academic use of cloud RI and commercial cloud systems. The barriers discussed included:

- Inertia regarding change to existing practices. In some cases, it's not about users making a choice, but it's not realizing that here **is** a choice, or even an issue to consider;
- Lack of knowledge or training; the "lack of training as experimentalists" that may require new or different ways of thinking;
- Fear of over-spending, whether this is a real concern or a misperception;
- Short-term cost or burden of moving existing data, code or workflows into the cloud;
- Micro-metered data egress cost a significant barrier; researchers express a strong preference for the predictability of a non-metered flat fee;
- Lack of 10,100,1000 Gb/s access ports; cloud providers thought to be unwilling/unable to provide this capability;
- Steep platform learning curve; users forced to 'grok' a lot of information to move forward; and
- fear of how long data will remain accessible.

But the group noted that benefits that would be enjoyed by overcoming these barriers include:

- Cloud skills developed by students would be useful for their careers; and the
- Ability to build on existing frameworks and avoid roll-your-own solutions could be a powerful research accelerator.

The group also noted that approaches to lower barriers would include:

Better documentation, and sample code/templates beyond just documentation;

Creating multiple `entry points' for users with different levels of experience and knowledge of cloud technologies (i.e., "You must be this tall to ride").There is a need to strike a balance between making it easy and requiring the user to have a certain level of understanding.
But there is a tension here. For example, CloudLab tries not to abstract-away low-level details. The challenge is to reconcile high-level views with lower-level views.

The creation of cloud "consultancy groups" or centers of expertise, and aa `help desk' to get basic code working;

Teaching users that cloud platforms have a very complex interface and a big attack surface. The creation of software and templates could enforce best practices.

Tools to help users select the right instance type, base machine images, software libraries, etc. This includes better tools to determine if any particular VM is idle and “forgotten” or is doing useful work, as unsophisticated users can leave idle resources in place and run up costs unnecessarily.

Support of new organizational constructs and team roles where domain scientists are paired with cloud research software engineers.

Develop new methods of creating onramps for people to use new tools, such as with teaching scripts.

Spread awareness of useful tools. For example, Zenodo handle some hard problems in storage, metadata versioning, etc.

Spread awareness of successful projects and reasons for their success. Pangeo was a success in part because Jupyter Notebooks provided a high-level, familiar interface to geoscience users. Users are manipulating large datasets, auto-scaling using DASK, but the user sees a web page where they work with HTML, Markdown, Python. All three of these are very familiar to users, so the complexity of the Cloud and complex interfaces are hidden.

Breakout Session 2

a. Education & Reaching Underserved Institutions

Participants: Rick Mcgeer, KC Wang, Mark Cheung, Deep Medhi, Jack Brassil

The group agreed that access to CRI is especially critical for under-represented institutions, and that larger institutions can more easily get commercial systems, such as the use of Heroku for scalable computing at UC Berkeley. The group noted that Minority-Serving Institutions (MSI) face challenges including:

- faculty occupied with heavy teaching load;
- relatively low sponsor funding;
- limited IT support even for basic administration; and
- limited other support resources on campus.

Given those constraints, the group suggested that key underserved institutions’ needs would include

- adequate network bandwidth for experimenters to use remote facilities or testbeds effectively, and
- deep integration of a testbed or facility into an existing teaching need. This teaching support would require

    - testbed support of resource reservations at specific scheduled times (e.g., class or lab meetings);
    - online help specifically for students; and
    - pre-build materials (e.g., learning modules) to support the specific MSI's existing curriculum.

The participants had constructive suggestions for NSF and research community support, including:

- programs to support onboarding institutions into CRI systems as done for the "NSF Cloud for All" workshop project;

- collaborations with private institutions including Schmidt and Mozilla foundations that have strong interest in supporting STEM education activities;

- supporting citizen science models (e.g., zooniverse.org) to engage students to perform tasks that benefit science while picking up critical skills; and

- NSF curriculum development grants directly targeting partnerships between underserved and minority-serving institutions and CRI projects such as the "Cloud in a box" project.

b. Goals for Future Cloud RI Systems – What Must be Built to Handle Experiments in 5-10 Years?

Participants: Paul Ruth, Glenn Ricart, Justin Shi, James Griffioen, Kate Keahey, Larry Landweber, Michael Zink, Anirban Mandal

The attendees discussed a diverse collection of goals for future CRI including:

1. Handling Protected Data

Research infrastructure must be capable of handling more types of real data to be more broadly relevant to more researchers. This ideally includes various classes of protected data such as Personally Identifiable Information (PII), Medical data (e.g., HIPPA), proprietary corporate information and algorithms, etc. In addition, some organizations like the National Portrait Gallery have very detailed image scans which not sensitive won't release for fear of counterfeiting or other data misuses.

2. Redefining cloud as a federation of tiered platforms

Ideally a future CRI would be a federated collection of heterogeneous computing systems supporting a high degree of experiment dynamics (e.g., migration of tasks between CRIs, execution placement between edge, fog and core cloud systems, etc). Ideally, an experimenter would be able to match a cloud's offering to their application need, as in selecting a system that provides the desired access latency.

3. Creating climate-neutral clouds

Energy consumption has not been a CRI priority, though its priority must increase as CRIs scale.

4. Merging research clouds and production clouds

Better ways to integrate research testbeds and commercial clouds will cause existing distinctions to diminish.

5. Driving future academic publishing with interactive papers and notebooks

Future CRIs should support repeatability through improved integration and support of online science publishing techniques and tools including Jupyter notebooks, Mathematica, etc.

6. Scaling

It will clearly be essential to build testbeds several orders of magnitude larger and more adept at storing, moving and processing massive data.

7. Enhancing testbed reliability, performance, and usability

8. Improving cloud transparency and understanding

New methods for supporting a more advanced ecosystem of tools to help experimenters use productive use CRI productively including API transparency.

c. Hardware Infrastructure

Participants: Chip Elliot, Joe Mambretti, Steve Schwab, Rob Ricci, Manu Gosain, Peter Desnoyers, DK Panda

Future cloud research infrastructures will face demanding experimenter requirements for both commodity and specialized equipment. The convergence of radio frequency access technology and edge cloud computing capability will guide future infrastructure. It is also anticipated that CRI will move beyond the conventional “racks of servers” implementation and introduce disaggregated compute/storage systems, direct-to-core optical communications, and considerably higher transmission rated in and out of the CRI facility. There will also be needed to quickly introduce anticipated new and revolutionary storage system technologies including NVMe and programmable storage.

It was noted that introducing newer technologies quickly requires CRI developers to engage in vendor partnerships as early as possible. There could be a role in which one or more open consortium assist CRI developers with the task of defining community equipment needs and equipment specifications and provide assistance in justifying the cost of acquiring resources.

The participants noted that research on CRIs has two goals, both advancing understanding (i.e., a conventional academic role) and helping to advance ideas that can shape the future of the commercial cloud industry.

To increase industry participation effort must be expended on consideration (and communication) of how a vendor benefits from the experience and participation of partnering with a CRI. It was noted that grad students play a crucial role in these partnerships.

A discussion focused on a desire to increase the level of risk in future CRI testbeds. While it was noted that demand will increase for specialized hardware such as GPUs and FPGAs, there is a need to avoid outcomes such as becoming GPU-farms for ML training applications. Some risk would be in the form of investing in exploratory technologies such as Quantum Processors (simulated), DNA storage, or Quantum Networks; another risk opportunity would involve making a bet on at-scale storage systems.

The participants agreed that providing bare metal service (and the value add of an OS on bare metal with installation of libraries, software, etc) is highly demanded and crucial for the research community.

The group also discussed topics that need additional analysis. For example, what is the optimal lifetime of testbed hardware? How critical is the need for supporting heterogenous system/processor types? How responsive should CRI developers and funder to support the relentless call for “more” (of existing resources)? How crucial is it for developers to expose low level interfaces and controls (e.g., ASICs) to experimenters?

d. Barriers to Academic Use of Commercial Cloud Systems

Participants: Jeff de La Beaujardière, Eric Eide, Amr El Abbadi, David Hancock, Jamie Sunderland, Kevin Tyle

The group recounted some of the barriers to academic use of commercial cloud from earlier discussions, including research inertia, lack of researcher training, and fear of absolute cost and costs overruns. Among the cited concerns associated with commercial cloud cost management were:

- the difficulty of allocating credits or costs among multiple users at institution;
- the challenge managing allocations and quotas (including hard limits) for students/projects, including high visibility into individual experimenter usage;
- local entity procurement funding limitations such as operational vs capital expenses;
- lack of collective buying power even across multiple academic institutions. The Internet2 Box agreement terms were noted as becoming less favorable after large buy-in; and
- academic institutions are reluctant to change practices after making Local investments in buildings, data infrastructure, etc.

The group observed that hybrid cloud approaches may be increasingly viable, and cited examples of 1) Big Weather Web moving from AWS to Jetstream; 2) Black Hole Imaging moving from Jetstream to GCP; and 3) Princeton neuroscience workflow between GCP and campus-based research computing storage facilities. Of course, certain types of research are not well suited for commercial cloud environments, including some climate modeling, security research, and networking experiments.

In certain cases, experimenters might be restricted from cloud use by policies such as data locality constraints (e.g., EU entity compliance with GDPR). While specialized data agreements are possible in some cases (e.g., a Business Associate Agreement for ePHI data covered by HIPAA), this can come with increased costs and large institutional startup cost to implement.

Technical difficulties associated with cloud migration were also note. One example is the translation of POSIX to Object storage semantics and interfaces, which are difficult for many domain researchers.

Among positive trends the group noted that commercial cloud providers have shown increasing interest in hosting data for commercially viable public/private partnerships (e.g. hosting NCAR data where industry wants to use for purposes not permitted on NSF-funded systems). It was also noted that commercial cloud use exposes students to popular cloud offerings, which is a win for workforce readiness.

The group expressed some concern that despite efforts like CloudBank that traditional funding mechanisms are lagging behind and have not caught up.
Recommendation: The participants expressed the desire for per-student “cloud credit gift cards” with hard spend limits but noted that providers have been reluctant to implement this payment mechanism.

Breakout Session 3

a. All Things Data

Participants: Kate Keahey, et al

Making metadata available about CRI operation is essential to enable research and reason about testbeds, and this is something that researchers want to analyze, yet industry does not and will not provide it. Both Chameleon and CloudLab capture data but make it available to researchers differently. While such data needs to be anonymized, it is unclear what are the data protection policies that need to be applied.

The participants found that each of the following need additional study or development:

- tools for data extraction, formats for representing the data, and tools for at least some basic data processing;
- identification and measurement of parameters of research interest (requests, power data, performance variability), and an assessment of obstacles for CRI developers to capture the data. The Chameleon experience suggests that not all desired information is available from log files. In some cases, benchmarks also need to be developed, and ways to publish the data (e.g. Zenodo) should be identified; and
- Recommendation: The best way to capture data is to build it in early in the process. Perhaps infrastructure data handling should be a required element of future CRI solicitations and construction proposals.

Real-time data about testbeds is of considerable research interest as evidenced by:

1. Chameleon records experiment setup (OpenStack) events that users performed on the testbed, such as creating leases, creating instances, and setting up networks. Users can request a report on their experiment records (*Experiment Precis)* using their Chameleon credentials. A report on those experiment records is known as the *Experiment Precis*.

2. Fabric plans to provide real-time information about testbed events. NetFlow data could be interesting to Fabric experimenters.

Making this information available from an API is desirable.

b. Federation

Participants: Paul Ruth, Glenn Ricart, Chris Hill, Justin Shih, Timur Friedman, Mosharaf Chowdhury

The federation breakout reviewed many of the difficulties that face ordinary researchers attempting to federate resources, whether their own with NSF resources, university resources with Cloud Access resources, or any other combination. We found no universal panacea.

The group, however, did spend some time thinking about what would need to be understood by resources being federated with each other. This isn't a solution to the federation problem, but rather a discussion of what would need to be involved in any solution that would make federation work.

Identity

Identity federation is well established in today's infrastructures. An open question is whether it is appropriate to suggest that each federated infrastructure should support a common identity provider (e.g., InCommon).

Namespaces

The first main concept was that the parties federating would need to be able to name themselves and the other party(-ies) to the federation. That in turn would imply an understanding of the union of the namespaces involved. Note that we are not assuming an eventual federation mechanism would necessarily agree upon a single namespace (though that would be convenient), but rather a way to describe the namespaces to be used. We noted in the breakout that today's Internet DNSSEC allows multiple character codes, national and international governance on top level identifiers, and a security mechanism to prevent masquerading and malicious DNS lookups. So, one might consider it as a root for federating namespaces.

Any given resource or service could have a unique name in multiple namespaces, but still be the same resource of service. This suggested that resources and services might have a "wallet" of identity cards in different namespaces. To the extent federating resources and services could agree on a single namespace, the job of managing the wallets would be simplified.

Methods/APIs/Services

Given that one can name another resource or service, the next thing a federation would seem to need is a list of ways in which resources or services can be invoked / called / linked to each other. This would seem to require describing eligible methods of invocation (APIs, RPCs, http requests, etc.)

Larry Peterson's work was mentioned as a possible starting point, and of course there are many other examples.

Credentials

In practical implementations, various resources or services may only be available to certain groups or to certified users, members of specific classes at educational institutions, etc. So a generic concept of a wallet of credentials would seem to be another reason for wallets. In addition to spanning namespaces, the wallet might have additional certificates for identify, authorization, tokens of value (or even actual money) for cases like commercial cloud resources where there needs to be a value transfer and/or limit on resources used, etc.

Management/Monitoring Interface/Channel

Any researcher experiment or student project assembled using federation should have a way of naming, managing, and monitoring the assembled federation. If for no other reason, every federated assembly of resources probably needs an identifier to enable an “emergency stop” capability.

Federation implies assembling useful graphs of resources and services along the lines of data planes and control planes (and possibly other planes) but at all times the owner of the federation needs to be able to manage it. The managers of individual resources may also need to send messages to the owner of a federation explaining issues such as “going down for maintenance” or “newer version available,” or “you are going to be preempted.” Such messages might begin with well-understood numeric codes followed by more specific information. If a caller did not have a specific response for a management plane message, it might understand a numeric code representing “temporary failure; feel free to try again” even if it didn’t understand the specific detailed failure information.

The management channel might be a two-way channel in the spirit of the Unix standard debug output which appears on a Unix console, or maybe something with more structure and complexity behind it.

Project and Instance Identifiers

To aid the use of federations, a configured federation might have its own name in a catalog. The catalog entry would specify some or all of the (generic) resources needed and the bindings (federations) between them. A student could then invoke a specific project without having to understand all of the details of the resources being federated on her or his account. The student would be given an instance identifier (or slice identifier) for their own copy of the project.

Of course, this is exactly what the role of advertising RSPECS and instance RSPECS played in GENI or “profiles” in CloudLab and PAWR.

Existing Federation Components

During the discussion, participants pointed out numerous examples of existing infrastructure components that could be drawn upon to assemble a federation architecture for research infrastructure. An incomplete list would include: CILogin, Globus identifiers, and RSPECs.

Extensibility

All of these concepts need to be able to be extended by individual research groups.

Binding GUIDs

To avoid the verbosity of specifying namespaces and name each time and in each interaction, we discussed binding a GUID to a verified namespace and name.

Key-Value Stores

Several people in the group noted repeatedly that *key:value* stores would seem to be a good underlying mechanism to express capabilities, list services, provide for expression in multiple namespaces, etc.

Uses

Any “Bring Your Own” research infrastructure (like FABRIC, which was represented in the group) will need to have mechanisms to extend to and link to the newly added capabilities. That raises the possibility of using “capabilities” as a first class object.

Recommendation: Federation is well suited to a become a major technical focus area for the Midscale Experimental Research Infrastructure Forum (MERIF), which provides a research network role spanning CISE midscale infrastructures.

c. Future Infrastructure Needs

Participants: Brent Stephens, et al

Finding: CRI needs to support specialized infrastructure to meet the needs of the systems researchers.

Systems researchers have very different needs than the science community. For example, programmable networks are not needed for bioinformatics today, although they are needed for some state-of-the-art systems research. However, advances in systems may be needed as part of a toolchain or framework

Finding: Future CRI needs software frameworks and open toolchains.

Exposing hardware in higher level packages to users is a key enabler of using new technology. For example, most users do not understand RDMA, but they are able to use systems and libraries that underneath the hood provide use these devices.

Finding: The research community would benefit from a more open process for directing future hardware acquisitions.

The current "whisper" process is not transparent enough and it leads to some people being unfairly privileged. A community process where a batch of hardware that is being considered where people can vote on the hardware that they plan on using is appealing. Also, this should be done earlier in the funding process, such as during the solicitation phase.

Opportunity: To address the problems caused by export controls attached to toolchains, there is potential for the academic cloud community to push back and bargain against these restrictions.

Open tool chains for programmable hardware are problematic because of export controls. In some cases, the research community could just pay to redevelop the small part of the toolchain that is controlled. In addition, we should be pushing as a community as much as possible to avoid closed toolchains.

The participants found the term *programmability* is too generic; the type of programmability being provided needs to be made more explicitly clear. For example, control and data plane programmability are very different types of programmability.

Dynamic allocation of VLANs is an example of a problem that needs tooling. For example, these VLANs could be created by hand by operators, but this requires too much effort. Experimental clouds need to be more like public clouds with observability into experiments. Users need better job statistics to be immediately shown to them. Programmability is likely a key enabler of this.

Complexity in infrastructure is a problem when different users are running experiments. Support and ticketing systems are a necessary part of supporting users who are going to have difficulty with using the technology. Can we embed it into a library so as to isolate small community of developers to create tools for the larger community to use?

It is important to make a distinction between hardware and software needs. Most advanced hardware capabilities need testbeds that do not pretend that they are going to serve broader science needs today. There need to be special tracts for incremental deployment. Service provisioning and scheduling is a key infrastructure requirement. People who are creating frameworks and APIs are key to enabling eventual adoption by other communities.

Service provisioning and scheduling is a key infrastructure requirement. It's important for new technologies to be available for experimentation and not tied up with allocations for the more general crowd.

There is a tension between federation and providing raw hardware with new capabilities. The general philosophy that is currently best practices is building as much software services as possible but not requiring that the full software stack be used and allowing users to bring their own stack.

d. CRI Measurement, Monitoring, and Experiment Debugging

Participants: Eric Eide, Rob Fatland, Jim Griffioen, Irene Kopaliani, Stephen Schwab

Measurement and monitoring of individual experiments is challenging since each experiment is complicated and multi-layered. Experimenter intent is opaque to CRI operators.
The group pointed to the example of GENI. A considerable amount of experiment measurement and instrumentation was built (e.g., dashboard, packet counters) and these were used quite heavily. But operator tools were less used, though it was unclear why. One possible explanation was that experimenters would call the help desk rather than debug infrastructure problems themselves. In this case, the experimenter takes the position "I just want it to work."
The group engaged in discussion to identify underlying sources of problems. It was noted that in some cases a GENI problem report exposed a misunderstanding of some experiment abstraction, e.g., using physical node names rather than logical per-experiment aliases, or confusion between the control and experiment networks).

Opportunity: Support more "campus champions" to help local experimenters use CRI.

Among potential reasons experimenters struggle are 1) lack of knowledge of pitfalls of the platform; 2) lack of platform documentation or training; 3) lack of knowledge about available platform experiment support tools; or 3) lack of training in the domain of their experiments.
Opportunities for helping experiments include 1) actively monitoring setup of experiments and alerting for failures, e.g., when VLAN stitching fails; 2) providing telemetry data to help desks; 3) actively running tutorials to make sure they are up to date. iMinds does this for GENI; and 4) providing tools that users can run to check experiment properties.

Two examples of such tools are:
1. Chameleon's *experiment precis*; and

2. the cloud troubleshooting tool CloudSight. See Hyunwook Baek et al., “CloudSight: A tenant-oriented transparency framework for cross-layer cloud troubleshooting,” CCGrid ‘17.

References

[cloudlab] R. Ricci, et al, “Introducing CloudLab: Scientific infrastructure for advancing cloud architectures and applications," *USENIX ;login:*,vol. 39(6), Dec. 2014.

[chameleon] "ChameleonCloud", https://www.chameleoncloud.org.

[jetstream] "JetSteam", https://jetstream-cloud.org/.

[vmware2] NSF 18-540, "NSF/VMware Partnership on Edge Computing Data Infrastructure (ECDI)", https://www.nsf.gov/pubs/2018/nsf18540/nsf18540.htm, 2018.

[ppo] NSF 16-585, "Platforms for Advanced Wireless Research (PAWR): Establishing the PAWR Project Office (PPO) (PAWR/PPO)", https://www.nsf.gov/pubs/2016/nsf16585/nsf16585.htm, 2016.

[dcl] NSF 19-013, "Dear Colleague Letter: Mid-Scale Research Infrastructure (Mid-scale RI) Opportunities", https://www.nsf.gov/pubs/2019/nsf19013/nsf19013.jsp, 2018.

[msri1] NSF 19-537, "Mid-scale Research Infrastructure-1 (Mid-scale RI-1)", https://www.nsf.gov/pubs/2019/nsf19537/nsf19537.htm, 2018.

[ecas] NSF Award \#1904444, "Exploring Clouds for Acceleration of Science (E-CAS)", https://www.nsf.gov/awardsearch/showAward?AWD_ID=1904444\&HistoricalAwards=false, 2018.

[bigdata] I.A.T. Hashem, et al, "The rise of Big Data on cloud computing: Review and open research issues," *Information Systems*, vol. 47, pp. 98-115, https://www.researchgate.net/publication/264624667_The_rise_of_Big_Data_on_cloud_computing_Review_and_open_research_issues, 2015.

[cloudaccess] NSF 19-510, "Enabling Access to Cloud Computing Resources for CISE Research and Education (Cloud Access)", https://www.nsf.gov/pubs/2019/nsf19510/nsf19510.htm, 2018.

[google] Yevgeniy Sverdlik, "Cloud Giants Continue Pouring Billions into Data Centers", *Data Center Knowledge*, https://www.datacenterknowledge.com/cloud/cloud-giants-continue-pouring-billions-data-centers, Oct. 26, 2018.

[bigstep] "BigStep," https://www.bigstep.com, 2018.

[onmetal] "RackSpace OnMetal," https://www.rackspace.com/en-us/cloud/servers/onmetal, 2018.

[tpu] “Cloud TPU,” https://cloud.google.com/tpu/.

[ec2-gpu] “Amazon EC2 elastic GPUs”, https://aws.amazon.com/ec2/Elastic-GPUs/.

[lambda] “Amazon Lambda," https://aws.amazon.com/lambda.

[xpu] M. Ferdman, A. Adileh, O. Kocberber, S. Volos, M. Alisafaee, D. Jevdjic, C. Kaynak, A. D. Popescu, A. Ailamaki, B. Falsafi, "A case for specialized processors for scale-out workloads," *IEEE Micro*, vol. 34(3), pp. 31-42, 2014.

[edgecloud] M. Chiang, W. Shi (Organizers), "NSF Workshop Report on Grand Challenges in Edge Computing," http://iot.eng.wayne.edu/edge/NSF%20Edge%20Workshop%20Report.pdf, October 26, 2016, Washington, DC.

[cloud3] A. Akella, G. Porter, K. Winstein (Organizers), "NSF Cloud 3.0 Workshop Report," http://pages.cs.wisc.edu/~akella/cloud3workshopreport.pdf, Stanford, CA, 2018.

[academiccloud] M. Balazinska, D. Culler, Jen Rexford, J. Wing, "NSF Workshop Enabling Computer and Information Science and Engineering Research and Education in the Cloud", https://dl.acm.org/citation.cfm?id=3233928, Jan. 8-9, 2018, Washington, DC.

[merif] NSF #1834880, "MERIF: A Forum for CISE-supported Midscale Experimental Research Infrastructure projects Projects", https://www.nsf.gov/awardsearch/showAward?AWD_ID=1834880\&HistoricalAwards=false, 2018.

[nationalacademies] The National Academies Press, "Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020", https://www.nap.edu/catalog/21886/future-directions-for-nsf-advanced-computing-infrastructure-in-2017-2020, 2016.

[cloudlab4all] "CloudLab for All Workshop", https://www.cloudlab.us/nsfcloud-for-everyone.php, 2017.

[cloudinbox] NSF #1842679, "EAGER: Cloud-Ecosystem in Box (CEB)", https://www.nsf.gov/awardsearch/showAward?AWD_ID=1842679, 2018.

[clusters] NSF 1800946, “Workshop on Clusters, Clouds, and Data Analytics in Scientific Computing", https://nsf.gov/awardsearch/showAward?AWD_ID=1800946, 2018.

[cord] L. Peterson, "Central Office Re-envisioned as a Data center," https://opencord.org/about/, 2019.

[osg] "Open Science Grid", https://opensciencegrid.org.

[opencloud] A. Bestavros and O. Krieger, "Toward an Open Cloud MarketPlace: Vision and First Steps," *IEEE Internet Computing*, January/February, 2014.

[krieger] O. Krieger, NSF #1744308, "NSF Cloud Evaluation Committee", https://www.nsf.gov/awardsearch/showAward?AWD_ID=1744308, 2017.

[nsfcloud] NSF 13-602, "CISE Research Infrastructure: Mid-Scale Infrastructure - NSFCloud (CRI: NSFCloud)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504951, Sept. 2013.

[bigdatahubs] NSF 18-598, "Big Data Regional Innovation Hubs (BD Hubs): Accelerating the Big Data Innovation Ecosystem", https://www.nsf.gov/pubs/2018/nsf18598/nsf18598.htm

[akella] A. Akella, "Experimenting with next-generation cloud architectures using CloudLab," *IEEE Internet Computing*, vol. 19(5), 2015.

[cloudlab-mgmt] D. Duplyakin, R. Ricci, "Introducing Configuration Management Capabilities into CloudLab Experiments", *Proceedings of the International Workshop on Computer and Network Experimental Research Using Testbeds* (CNERT), April 2016.

[hubs] NSF 18-598, "Big Data Regional Innovation Hubs (BD Hubs)", https://www.nsf.gov/pubs/2018/nsf18598/nsf18598.htm, 2018.

[emulab] F. Hermenier, Ricci R., "How to Build a Better Testbed: Lessons from a Decade of Network Experiments on Emulab",In: Korakis T., Zink M., Ott M. (Eds.), Testbeds and Research Infrastructure. Development of Networks and Communities, *TridentCom 2012*. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 44, Springer, Berlin, Heidelberg.

[future-aci] National Academies of Sciences, Engineering, and Medicine 2016. *Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020.* Washington, DC: The National Academies Press.

Appendices

Appendix A: Workshop Participants

Name	Institution
Jack Brassil	Princeton University
Mosharaf Chowdhury	University of Michigan
Kuang-Ching Wang	Clemson University
Justin Shi	Temple University
Mohammad Hajiesmaili	UMass Amherst
Robert Ricci	University of Utah
Joe Mambretti	Northwestern University
Abhimanyu Gosain	Northeastern University
Jamie Sunderland	Internet2
Tamar Eilam	IBM Research
Rob Fatland	University of Washington
Mohan Ramamurthy	Unidata/UCAR
Brent Stephens	University of Illinois at Chicago (UIC)
David Hancock	Indiana University
Dhabaleswar Panda	The Ohio State University
Brian Voss	Indiana University
Kevin Tyle	UAlbany-SUNY
Mark Cheung	Lockheed Martin Space, USA
Glenn Ricart	US Ignite and U Utah
Kate Keahey	University of Chicago
Paul Ruth	RENCI - UNC Chapel Hill
Peter Desnoyers	Northeastern University, Boston University
Michael Zink	UMass Amhert, BU, Northeastern
Orran Krieger	Boston University/MOC
Amr El Abbadi	University of California, Santa Barbara
Rick McGeer	US Ignite
Ryan Abernathey	Columbia University
Joe Hamman	NCAR
Jeff de La Beaujardière	NCAR
Chris Hill	MIT
James Griffioen	University of Kentucky
Chip Elliott	Raytheon BBN Technologies
Jennifer Rexford	Princeton University
Eric Eide	University of Utah
Stephen Schwab	USC/ISI
Sebastian Seung	Princeton University
Larry Landweber	MERIF

Li-Fen Wu	Microsoft
Erica Lan	Microsoft
Anirban Mandal	RENCI - UNC Chapel Hill
Dan Stanzione	TACC/Univ. Of Texas
Will Silversmith	Princeton University
Deep Medhi	NSF
Ann Von Lehmen	NSF
Timur Friedman	Sorbonne
Irene Kopaliani	Princeton University

Appendix B: Online Community Survey of Cloud Research Infrastructure

Objective

The primary purpose of the online web survey was to increase research community understanding of how existing CRIs are being used for experimentation today, and to identify researchers' existing and emerging needs for CRI support of future experimentation. Secondary purposes included exploring

1. current and future use of CRI for educational uses (e.g., laboratory projects, homework assignments, in-class use, etc.); and
2. current research community use of related and/or complementary computing facilities including commercial cloud systems, university-hosted shared research computing systems, and related (non-cloud oriented) computing and networking systems research infrastructures (e.g., GENI, PAWR).

Survey Development

Tool Selection

Anecdotal evidence gather from colleagues who published prior community surveys suggested that 1) low survey response rates should be anticipated, and 2) direct personalized outreach would be more effective in evoking responses than bulk email requests for participation. Anticipating modest needs for online web survey capabilities and post processing analytics, we elected to use a simple Google Forms survey rather than a more general-purpose tool such as Qualtrics XM Survey Tool.

Process

An initial draft survey was created by the PIs with some NSF Program Director input. This draft was circulated to and reviewed by several informed community members. A common concern reported by reviewers was the tension between the desire to collect detailed information, while not expending too much respondent time and effort to complete the survey. At risk of introducing bias, we decided to solicit review input form the PIs of the *CloudLab* and *Chameleon Cloud* projects; we believe the survey would benefit from their considerable understanding of how existing platforms are used, and their considerable experience working with their user communities. We also believed that the obtaining existing CRI platform PI buy-in for the survey would potentially help with any subsequent community outreach (if needed) to increase initial survey response. In retrospect, involving the CRI PIs provided good initial feedback on the draft survey, and likely increased participation.

We elected to not personalize email requests for survey completion, though a limited degree of personalization would not have been technically difficult. We deferred taking this step, preserving it for a subsequent round of participation requests if needed. It ultimately was needed. Rather, we sent a generic email invitation to the 2350-person *cloudlab users* email list, and also the 2775-person distribution list for Chameleon experimenters with alias *users@chameleoncloud.org*. These lists contain both past and current users and are believed to nearly fully represent the vast majority of each CRI's experimenter community; we noted that less than 4% of the cloudlab emails bounced back due to outdated or incorrect addresses. An additional hundred or so research community members (some overlapping the email lists) were invited to attend the CRI workshop and also strongly encouraged to participate in the survey.

Bias

Given the community is fairly small and tight knit, we entered the process with no expectation that the survey results would be unbiased. On the other hand, by inviting some feedback as short text responses we were confident that experimenters would freely present unvarnished feedback. Where possible, we strove to use generic language over testbed-specific language to ensure a common understanding of technical terms. Where it was necessary to mention the names of commercial cloud product or service offerings, we adopted AWS product language (e.g. Simple Storage Service or S3) as AWS is has the largest footprint in higher education and these service labels would likely be broadly recognizable. The survey was publicly accessible; response might have come from interested parties outside the targeted email lists, and this input was welcome. We requested that responses contain the names and contact information to avoid duplicate submissions.

Survey Response

In this section we highlight important results obtained from user survey, and include some basic analysis of results. For a complete examination of survey results, we encourage the reader to visit the survey questions (https://tinyurl.com/unrp8ja) or via

Figure 1: Scan for access to survey questions

The survey response with interactive graphics is best viewed online (https://tinyurl.com/rdhn78f) and can be reached at

Figure 2: Scan for access to survey results.

Experimenter Community & Experience

Faculty, students, and researchers represented over 90% of the 212 survey respondents. The fraction who have previously performed some CRI experimentation identified use of CloudLab (69.8%), Chameleon Cloud (47.2%), and JetStream (8%). While it is unsurprising that some experimenters have used multiple infrastructures over time, CloudLab and Chameleon Cloud are used as the predominant infrastructure by 59.4% and 31.6% of experimenters.

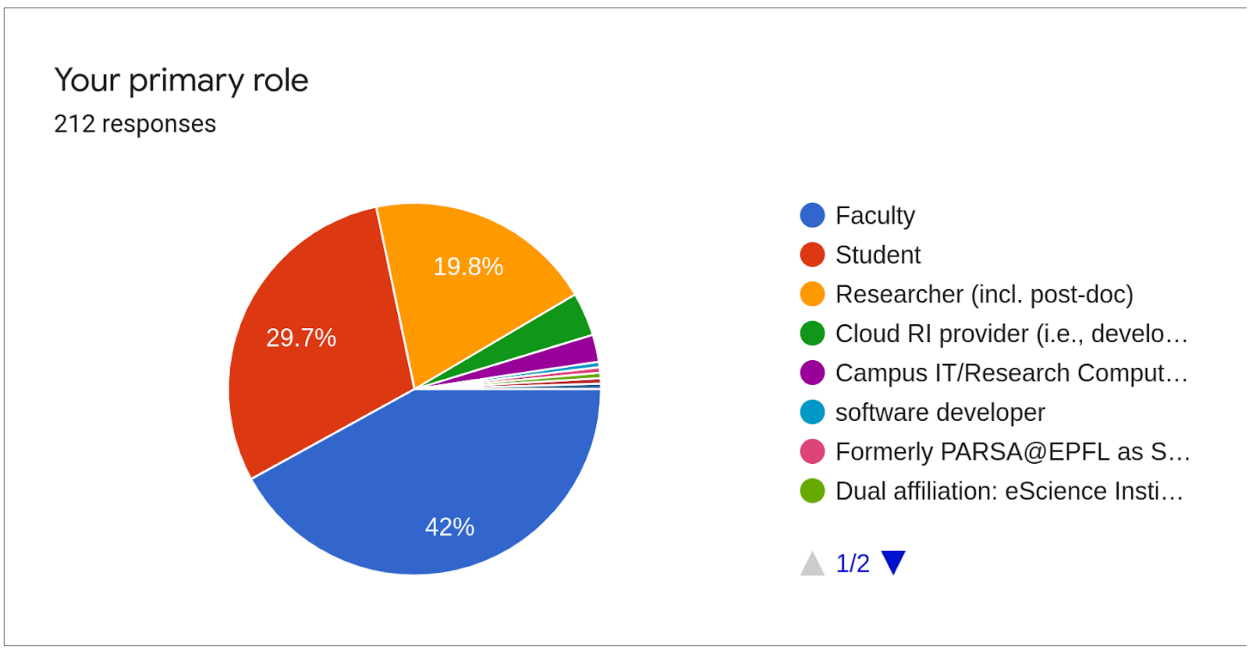

Figure 3: More than 2/3 of survey respondents were faculty and academic professionals.

Overall, the experimenter community reports being very satisfied with the existing CRI systems, with nearly 85% of responses rating satisfaction as 4/5 or 5/5 (5 is best).

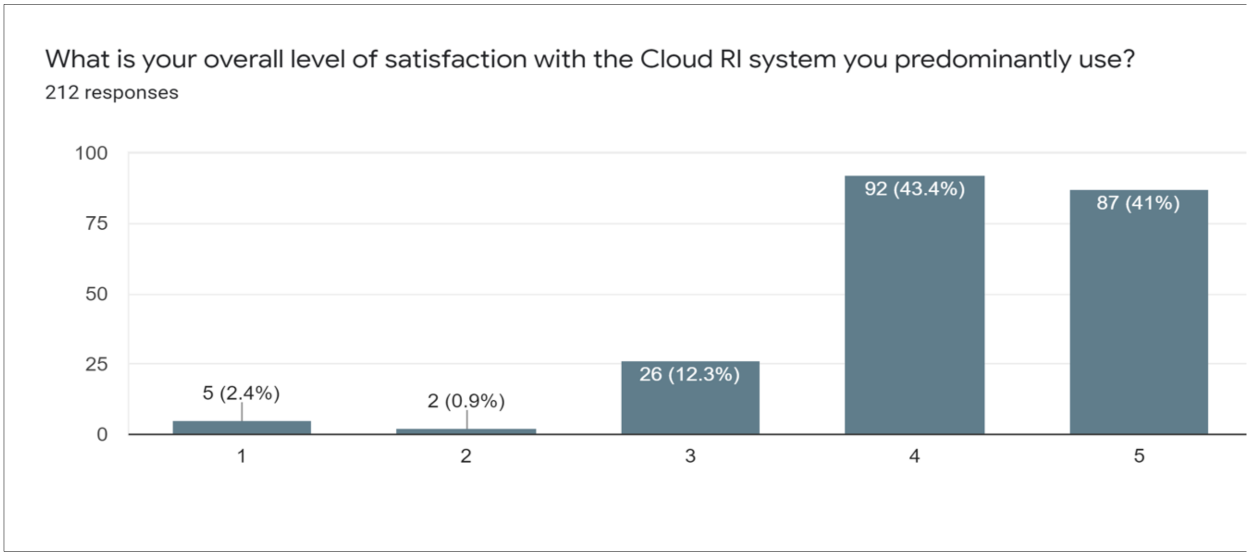

Figure 4: The research community reports being highly satisfied with available infrastructures.

Approximately 2/3 of the respondents learned how to use CRI online, bootstrapping themselves either through instructional documentation or online or hands-on tutorials. The bulk of the remainder learned CRI experimentation through instruction by a knowledgeable colleague. Experimenters also expressed satisfaction with these online resources, which seems at some odds with community concern that education and outreach are less of a priority than infrastructure replenishment and enhancement. The usability of infrastructures was also considered satisfactory, perhaps because of experimenter familiarity with the previous generations of NSF-funded RI projects – such as GENI and Emulab – that presented similar user experiences.

In terms of experiment creation, both standard and custom images are used significantly. Almost all users still use standard Linux tools to manage their configurations.

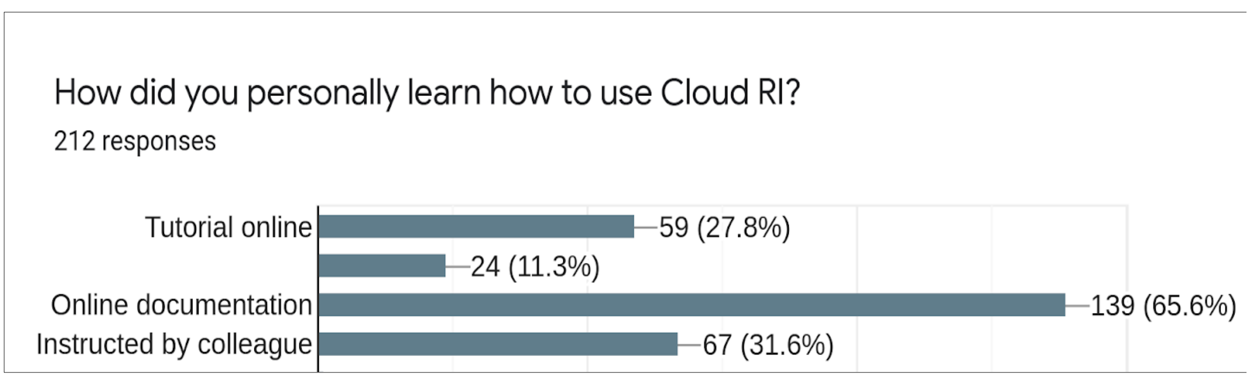

Figure 5: Online tutorials and documentation are the predominant means of learning how to experiment with CRIs.

Time expended using a CRI is considerable; nearly half of experimenters claim to hold CRI resources on average of roughly 10 hours each week, or more than one 8-hour working day.

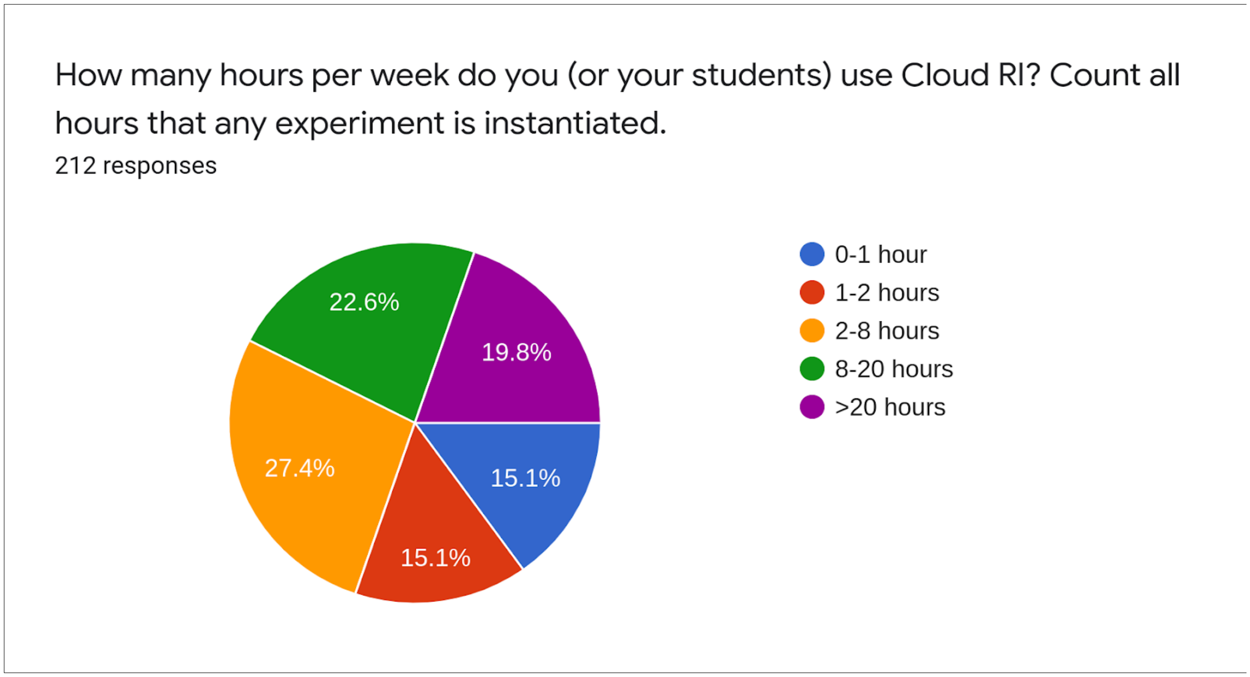

Figure 6: Active CRI experimenters invest considerable time in CRI experimentation.

The survey also examined the hypothesis that CRI functions as a meeting place for community building and information sharing. Slightly over 50% of respondents indicated a sense of belonging to a community, with a bias toward experimenting (and contributing) more over time than less over time. A surprisingly high one-sixth of respondents contribute actively to improving and maintaining the community resources.

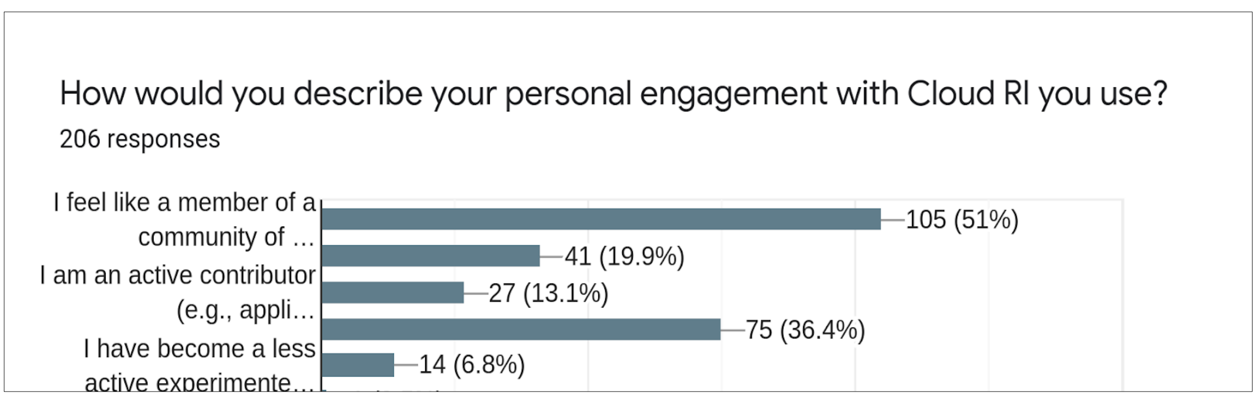

Figure 7: Survey respondents report a sense of belonging and contributing to a community of CRI experimenters and an associated knowledge base.

There is growing evidence that CRI is consequential in supporting both research and education. First, those surveyed reported research as the primary use of CRI resources. And two-thirds of the respondents have published one or more research papers in top conferences that include experimental results developed on CRI resources. Teaching graduate and undergraduate courses is the second largest CRI application. In this regard, the most common activities are research projects in graduate courses; some faculty experimenters use CRI resources for homework and lab assignments as well.

Experimentation

The survey explored how researchers use CRI resources for experimentation.

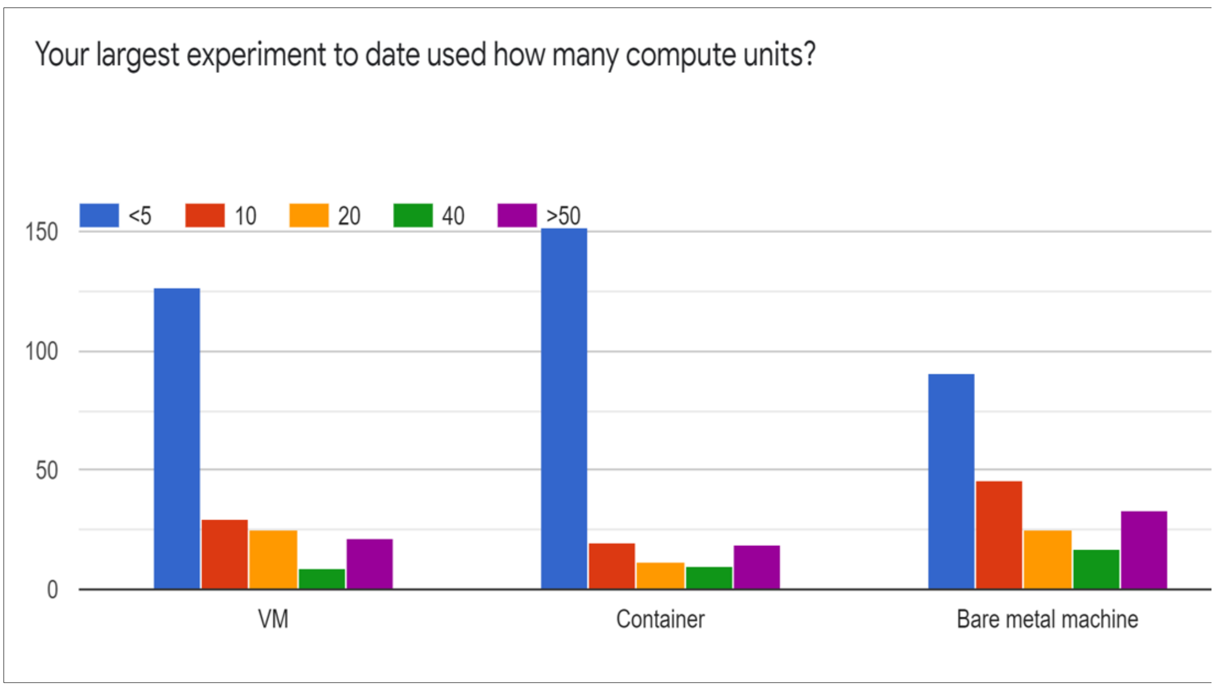

Figure 8: Even larger experiments remain modest resource consumers. Despite the presence of additional resources, experiment size does not appear to be growing materially over time.

Nearly a decade ago operators of Emulab [emulab] observed that typical experiments were small consumers of physical resources, and even larger experiments were modestly sized relative to CRI system size. It is intriguing to see – despite the much greater availability and ease-of-use of virtualization technologies such as containers and virtual machines (Vms), and the availability of more computing resources (servers or cores) -- that typical experiments remain small.

The availability of more hardware resources does not appear to increase the duration of the instantiation of an experiment; more than 3/4 of experiments operate for 6 months or less. We are left to speculate about the persistence of relatively short experiment durations. Short experiments might be tied to project cycles such as the time to publish a conference paper. Experiment duration might be closely tied to the research topic area or the need for external users to participate in the experiment. For example, *PlanetLab* hosted a variety of long-running experiments in research areas popular at that time, including caching systems, overlays, peer-to-peer networking, and content distribution systems; some of these systems provided service to third party clients over extended time periods. Of course, experiment duration might also reflect the growing ease of tearing down and re-instantiating experiments in modern CRI.

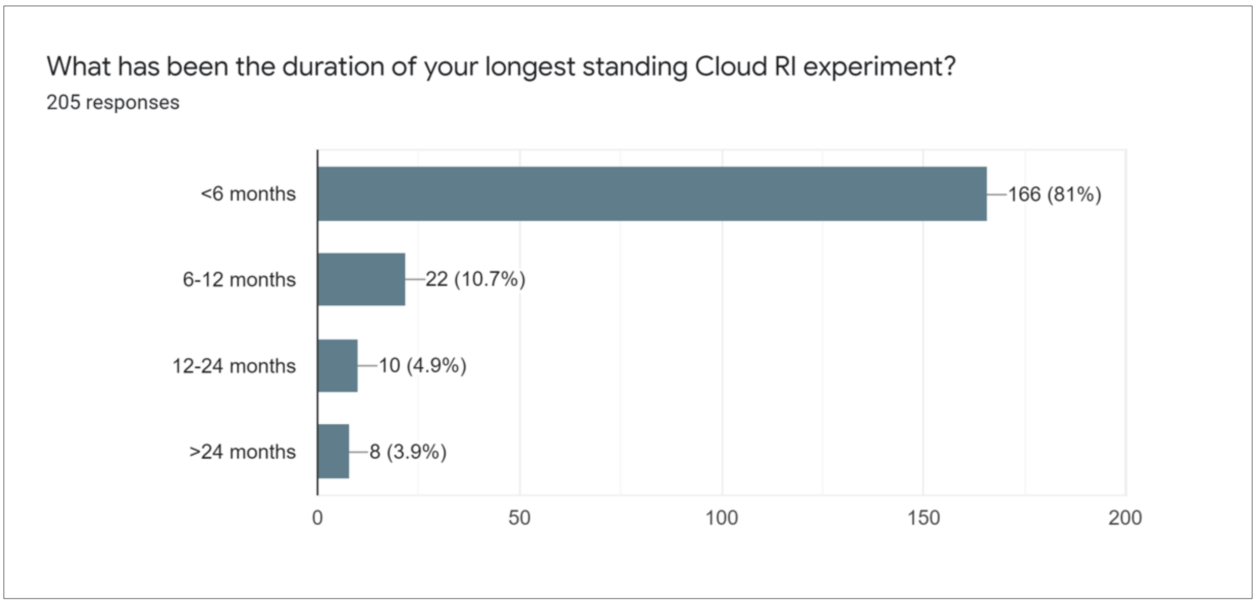

Figure 9: Even the longest running experiments rarely stay instantiated for a year or more.

Cloud RI today supports a wide variety of research projects, with networking, computing, and applications being the top three broad research areas. The diversity of research topics complicates anticipating which infrastructure investments that will be most desired in future experimentation. In terms of desired investment in specific hardware components, more CPUs remain the most preferred request of experimenters, though the need for GPUs, SSDs, and SDN is also increasing rapidly.

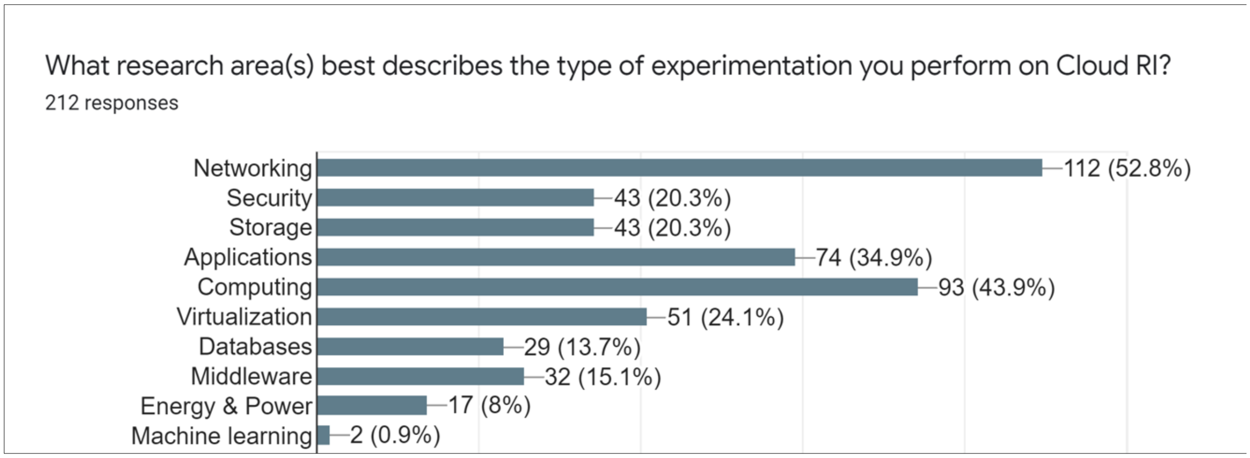

Figure 10: CRI supports the spectrum of computer and networking systems research topics.

Although most experimenters prefer on-demand resource allocation, a significant amount of them highlight the need for pre-allocated resources to have predictable resource availability. Given that most experiments are shorter than six months, both models have their pros and cons. Cloud RI users did not identify any significant barriers toward their usage of these resources. However, as expected, increasing resource availability can further improve usability. Overall, CRI resources are well utilized and operating at high efficiency.

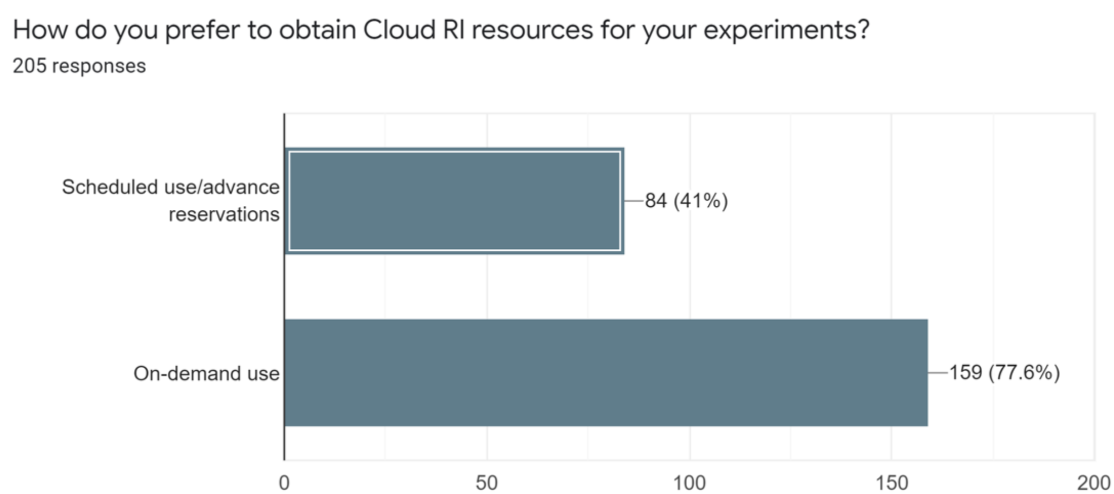

Figure 11: Experimenters seek both on-demand and scheduled access to CRI resources.

In terms of experiment development support tools, users appear of mixed minds. They make considerable use of both standard and custom disk images, profiles and appliances, all of which support replication and speed experiment construction. On the other hand, almost all users still use standard Linux tools for configuration management, though dedicated tools are available and often more feature rich. Additionally, experimenters suggested that improving exposure to internal details of the infrastructure would help users better understand their experimental outcomes.

Role of Commercial Cloud Computing

Experience using commercial cloud systems is increasing with responders. Unsurprisingly, AWS -- the most established provider in higher education – continues to be the system the highest percentage of responders have reported using.

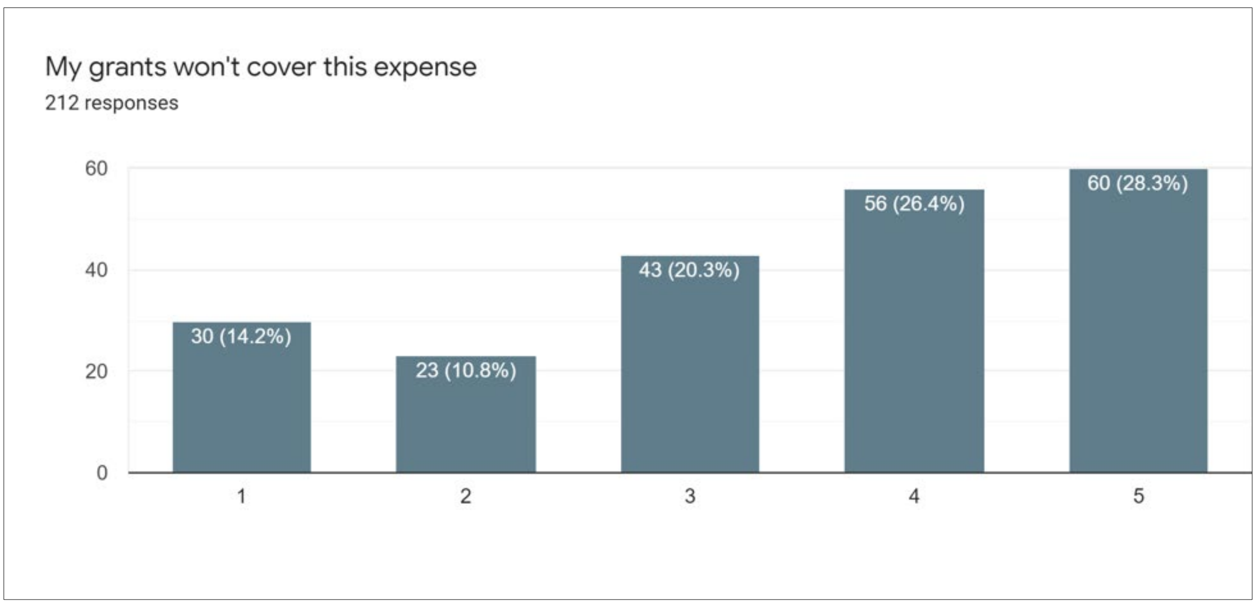

Figure 12: Experimenters with legacy grants cite lack of support for paying cloud computing service charges.

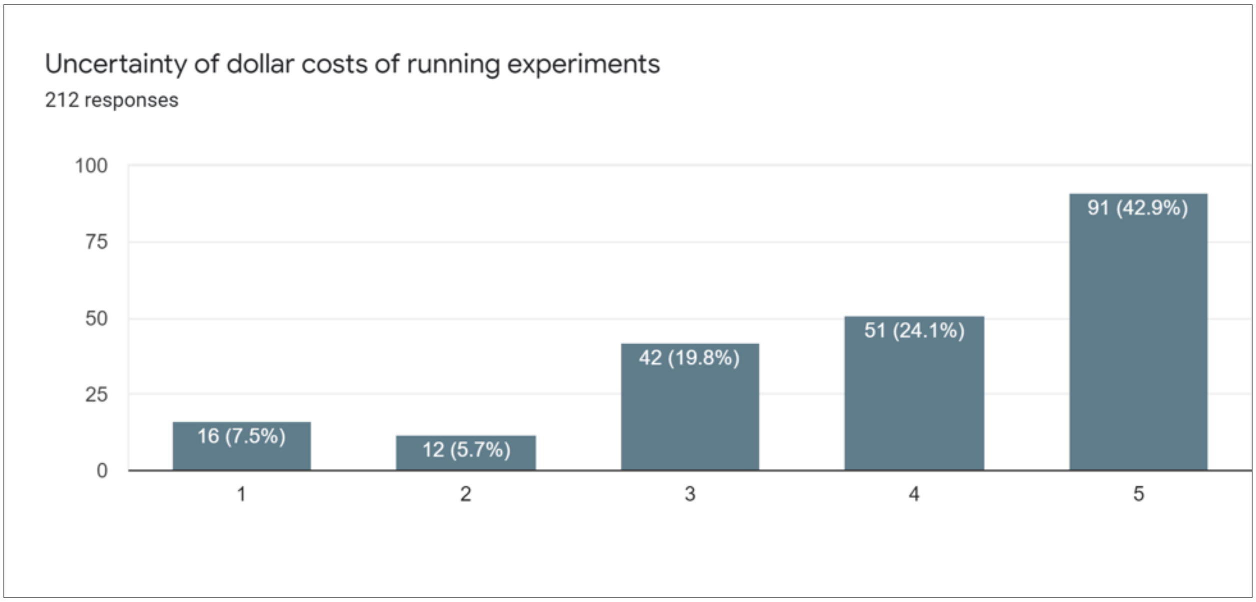

Figure 13: Most experimenters cite cost uncertainty as a significant barrier to experimentation on commercial cloud computing alternatives to CRI.

Most of the users prefer cloud RI resources because they are available and the least costly option today to carry out their research. They also use commercial clouds to complement their research requirements. However, such usage is limited by the cost of commercial clouds and the fact that most legacy grants do not support for commercial cloud expenses. A deeper

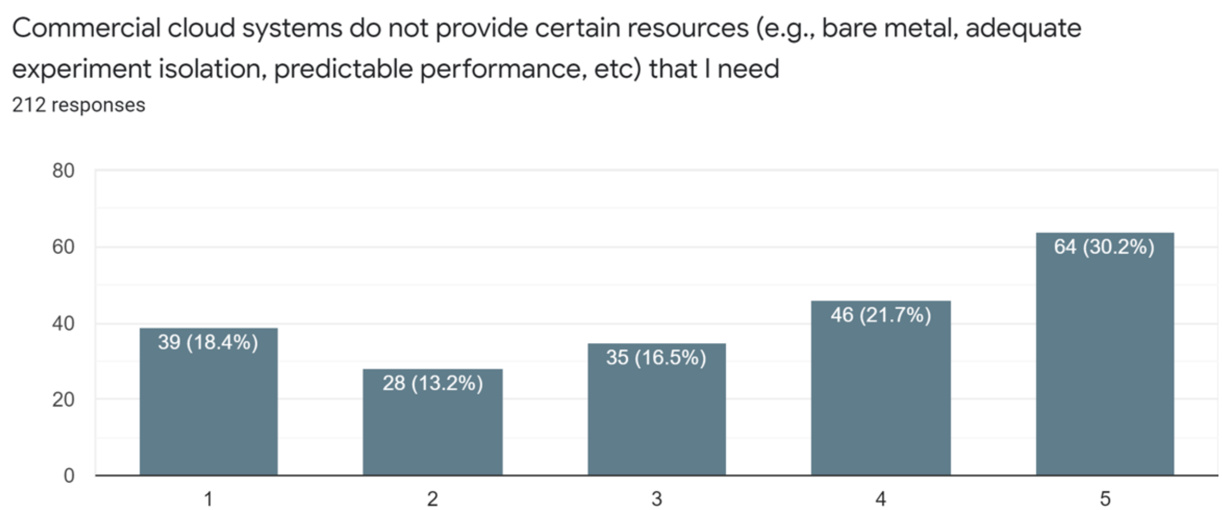

Figure 14: Commercial cloud services do not offer certain desired capabilities that can be found in some CRIs designed for science experimentation.

exploration of community concerns related to paying for commercial computing services can be found in [academiccloud]].

Conclusion

Overall, the survey results indicate that the CISE research community finds current CRI systems meets critical community infrastructure needs, are well run, and provides a unique and important offering to the research community.

Workshop Submissions

Advance submission of 1-page abstract on a proposed presentation topic was strongly encouraged for workshop participants. The following is the complete set of abstracts received. Not all submissions appeared at the workshop; the authoritative list of presentations can be found at the workshop web site (https://sites.google.com/view/workshop-on-cloud-cri/agenda).

Terra: A Framework for Optimizing Hybrid Multi-Cloud Analytics
Mosharaf Chowdhury

As cloud computing comes off age, storing data and running computation in hybrid multi-cloud environments is rapidly gaining popularity. However, data analytics across on- and off-premise sites connected by a patchwork of network links, i.e., hybrid multi-cloud analytics (HMCA), is still in its infancy. Despite a massive literature on big data analytics within the same datacenter and recent works on wide-area data analytics, existing solutions focus only on minimizing the amount of data transfers from inter-site shuffles over the wide-area network (WAN) but ignore how they are transferred.

We observe that the latter issue is equally important, if not more. Specifically, taking the inter-site topology into account while transferring data for HMCA queries can allow us to judiciously control their scheduling, path selection, and band- width allocation; which, in turn, can improve performance and increase WAN utilization. We observe that simultaneously performing all three in a scalable manner raises unique algorithmic and systems design challenges.

To this end, we present a two-pronged approach to de- sign Terra, which includes (i) a scalable, provably optimal algorithm for inter-site shuffle, which we use as a building block to perform single- and multi-query optimizations in offline and online settings; and (ii) a scalable enforcement mechanism that can quickly react to runtime uncertainties such as WAN bandwidth fluctuation and link failures. Integration with Apache YARN and evaluation using TPC-DS, TPCx-BB, and TPC-H workloads on three HMCA deployments shows that Terra improves HMCA performance by up to 3.43-32.04x w.r.t. the state-of-the-art.

Clouds and Internet Converged for the Future

Kuang-Ching Wang

To date, cloud computing has demonstrated how computing service can be delivered over the network at scale ‚Äì both in numbers of applications and distinct users - with elasticity, portability, customizability and cost effectiveness. The Internet that connects clouds to their stakeholders (service providers, end users, systems and devices), however, is operated in a model that poses a stark contrast. Internet has always been run with a significantly simplified core with as little customized handling as possible, in favor of higher volume transferred at lesser per-bit costs. For that reason, many of the important elements of today's applications, such as customization and security, have been done at the edge of the Internet. This has been lauded as a genius architectural principle that has sustained the fast and massive growth of Internet and networked applications so far. Looking into the future, however, there are many reasons for us to revisit this topic and envision a new Internet paradigm, where network services can be just as customizable and scalable to Internet-scale, diverse customers as cloud computing services can today. One path towards realizing this vision is by infusing into the next generation Internet the architectural and operational properties of cloud computing systems. By building the next generation Internet infrastructure like cloud system infrastructure, by allowing customers big and small purchase and provision customizable network services on and among such infrastructure, and by facilitating the deployment of software at flexible locations as needed to support specific application needs, the resulting paradigm is essentially a converged cloud and Internet. As a result, one can view a next generation cloud system a pervasively distributed system that can host applications at distributed locations and anywhere in the network path; one can also view the next generation Internet as a globally scoped cloud system to host distributed applications.

The proposed converged cloud and Internet paradigm has been discussed in various recent research efforts. From our NSF EAGER project on traffic analysis resistant networks, a globally distributed SDX architecture was introduced to invoke individually customizable security service for avoiding traffic analysis attacks throughout end-to-end application flows [1, 2]. In a community white paper on a future distributed network research infrastructure, the vision painted a future Internet that is programmable everywhere directly by applications [3]. For that to scale, one would expect whoever operates the next generation Internet to be capable of servicing such requests from a massive base, much like (but much larger than) today's cloud computing systems serving anyone purchasing computing instances with a credit card. Commercially, convergence of today's Internet service and cloud service's business models are expected. Academically, convergence of networking and cloud systems research are expected.

Lastly, it is my strong hope that future cloud research testbeds are not just another data center worth of computers. It is important that broad aspects of networking infrastructure and resources, such as 100 Gbps interfaces, BGP utilities, AS and IP addresses, interface to public Internet, public cloud style VM and/or container life cycle management tools, are provisioned as well. It is also important to facilitate recruitment and management of opt-in users in research experiments. It will be desirable to support global scale and inter-campus experiments, potentially leveraging available research IP ranges from international research network organizations and, potentially, U.S. universities and organizations.

References

1. K.-C. Wang, et al. Internet Security Liberated via Software Defined Exchanges. In Proceedings of ACM SDN-NFV Sec'18, https://doi.org/10.1145/3180465.3180475.
2. K.-C. Wang et al., NSF EAGER: Towards a Traffic Analysis Resistant Internet Architecture, https://nsf.gov/awardsearch/showAward?AWD_ID=1643020.
3. Ilya Baldin, et al. The Future of CISE Distributed Research Infrastructure. SIGCOMM Comput. Commun. Rev. 48, 2 (May 2018), 46-51. https://doi.org/10.1145/3213232.3213239.

Exploring new cloud services with NSF testbeds.

George Kesidis, Mahmut Kandemir, Bhuvan Urgaonkar

We're still in an era of diversified public cloud services, sometimes with poorly defined performance or resource guarantees in their SLAs and with diverse pricing rules. For different workloads, a fundamental question facing cloud customers is:

How best to use existing services at the lowest cost?

For cloud providers, a fundamental question is:

What services will be most profitable?

There is substantial research on reverse engineering cloud services (including our own) that informs the former problem. Testbeds can be used to verify such reverse engineering efforts.

Bare metal testbeds can also be used by researchers to prototype new services, study how useful they are in meeting the needs of different workloads, and study what pricing scheme would make them more attractive than existing services. Computer science education in cloud computing not only needs to address the performance gained by different service compositions, but also the costs.

Next Gen Cloud from the Science of Networked Computing ‚Äì A Proposal

Justin Y. Shi

Introduction: Science is a systematically organized body of knowledge on a particular subject. Computer science and network science are examples. For cloud computing, there is a need for Networked Computing Science. The lack of it has been demonstrated in cloud computing practices. Existing computing clouds are virtualization of physical infrastructures. The decoupling effects of hardware virtualization brought dramatic resource efficiency improvements. However, as the computing clouds expand, the probability of virtual machine crashes is increased due to resource sharing. All applications face the same scaling dilemma: employing more resources may improve application performance but suffers the increasing service downtimes and growing security attack vectors.

Networked Computing Science: There are two overarching theories: a) the impossibility of reliable communication in the face of crashes [1], and b) reliable system using statistic multiplexed unreliable resources [2]. The impossibility theory illustrates the need for application-level end-to-end communication [3] as opposed to the traditional hop-by-hop protocols. The reliable system construction theory proved the feasibility of highly reliable mission critical services overcoming the reliable failure detection difficulties. Unfortunately, existing distributed and parallel research and developments have been largely limited under the presumed reliable hop-by-hop protocols, such as message passing, remote procedure or method invocation methods and distributed memory sharing. These paradigms contributed to the scalability dilemma. Expanding computing clouds simply spread the dilemma in larger scales.

Proposed Research and Developments: Application-level end-to-end protocols for data intensive and compute intensive applications, especially for AI and machine learning applications and sensor network applications. These developments will enable complete data and program decoupling from physical resources without sacrificing performance, reliability or security. Consequently, research and developments of cloud resource optimization should focus on real time resource dispatching based on different application requirement metrics. Formal specification and validation of hard real time services become attainable.

Proposed Education and Training: The inclusion of Networked Computing Science theories will impact traditional theoretical CS courses especially for parallel performance/reliability analysis and prediction. Traditional CS programming training should formally introduce uncertainty and retransmission disciplines, operating system courses should include resource multiplexing in addition to existing cloud services, communication protocols should include BitTorrent and blockchain analysis and emulation, database courses should include CAP Theorem and its limitations, software engineering should introduce new resource optimization studies without application reliability concerns, formal validation and certification become practical, cyberinfrastructure courses should focus on end-to-end security and in-situ forensics and globally coordinated responses.

References:

[1]Alan Fekete, Nancy Lynch, Yishay Mansour, and John Spinelli. The impossibility of implementing reliable communication in the face of crashes. J. ACM, 40:1087-1107, November 1993.
[2]John von Neumann, Probabilistic Logics and Synthesis of Reliable Organisms from Unreliable Components, Automata Studies, p.43-98, 1956.
[3]J. H. Saltzer, D.P.Reed, D.D.Clark, End-to-end arguments in system design, online: http://web.mit.edu/Saltzer/www/publications/endtoend/endtoend.pdf, 1983.

Data Center Energy Optimization
Mohammad Hajiesmaili

Overview:
In recent years, there has been an unprecedented growth in energy footprint and cost of data centers as critical infrastructures of Internet services. Powering data centers from a portfolio of energy sources and elasticity in load execution are two key features that enable optimizing the energy cost by shifting the energy load in the temporal domain. This proposal intends to develop energy cost minimization algorithms that utilize the potentials of shifting the data center energy load in time. Having extreme and multidimensional uncertainty as the fundamental challenge, this proposal proposes a disciplined research based on algorithmic and learning understandings of the underlying optimization problems. Successful implementation of this proposal provides a beneficial design that can optimize the cost and robustness of data center energy operations and hence will have a significant impact on lowering the overall cost of Internet services.

Intellectual Merit:
This proposal focuses on three key goals of data center energy operations: minimizing the energy cost, maximizing the robustness against uncertainty, and improving the energy footprint of data centers. Toward this, it underpins the theoretical foundations of energy and load management in data centers from online algorithm and learning perspectives. Both are promising approaches since they do not rely on any exact or stochastic modeling of the future, hence, they are robust against extreme variability.
The cornerstone of this proposal is to decompose the general problem into two subproblems of energy procurement and load management. Then, it develops robust and cost-effective online algorithms for each subproblem with provable performance guarantees against severe uncertainty. Finally, by leveraging the insights from the algorithm design for the subproblems, and the optimal offline solutions, it develops efficient learning algorithms for the general problem. The intellectual merit of this proposal consists of the following:
1. Online Optimization for Energy Procurement: A disciplined design that tackles the cost minimization data center energy procurement. It designs online algorithms with provably best performance guarantees. 2. Online Optimization for Load Management: It designs cost-effective and deadline-aware algorithms for managing elastic loads in data centers, with provable competitiveness against optimal offline solutions.
3. Online Learning for Joint Energy and Load Management: A learning-based framework that learns joint optimization of the energy procurement and scheduling of the elastic load.
This research brings fundamental algorithmic and learning understandings into data center energy research and also extends the existing results on classic online conversion problems, i.e., the online search for best prices in order to buy and/or sell assets, with single-dimensional uncertainty, to the multidimensional uncertainty. Hence, the proposed research makes fundamental theoretical contributions by advancing conversion problems to address multidimensional uncertainty.

Broader Impacts:
This project has numerous avenues to create broader impacts. Data centers are key infrastructures that enable a variety of Internet services. Hence, our work on optimizing the cost of a data center will have a significant positive impact on lowering the cost of Internet services. More broadly, it will facilitate the efficient and reliable incorporation of renewables into the data center, thereby reducing the carbon footprint of data centers. Such improvements will play a key role in moving toward a more sustainable data center and the electric grid. The datasets and software packages of this project will be released publicly to the research community to help

with the comparative evaluation of their systems as well. Educational impacts include integrating modules on a new course under development by the PI on ""optimization under uncertainty in computer systems''. As part of diversity efforts, the PI will recruit women and under-represented groups to make a diverse research group.

Why We Need Testbeds for Cloud Computing

Robert Ricci

Getting resources in a cloud is easy - all it takes is a credit card (or some donated credits.) Given the easy availability of these resources, this begs the question: Why do we need to build testbeds for cloud computing - why don't we just use the plentiful, and relatively cheap, cloud resources that are available to us?

Of course, many research endeavors can ""just use"" the cloud, but there is an important set that cannot. This is the set of projects that want to improve the cloud itself, rather than the software that runs atop it. By its nature, the cloud presents layers of abstraction on top of the underlying hardware: the cloud provider manages the compute, network, and storage, and gives the user little control and no visibility into these layers. This abstraction is precisely what makes the cloud valuable: it's the cloud provider's control over these resources that enables it to make them available to users flexibly, at scale, and at low cost. The result is that the economic incentive of public clouds do not always align with the needs of the research community: many members of the cloud computing research community (in particular, those in the systems, networking, and security fields) need access to the system at a level different than what cloud providers offer. It is not the case that commercial providers cannot offer what these researchers need: it simply does not make business sense to do so.

This is why we need testbeds for cloud computing: so that researchers can help shape the future of cloud computing by getting access to resources that commercial clouds don't have incentives to provide.

This perspective suggests a strategy for prioritizing features for cloud computing testbeds: pursue features that enable new research into future cloud computing platforms, but that are not aligned with the business interests of cloud providers. This can take several forms. One is to look forward at technologies that are not yet available in commercial clouds, but which have potential to be in the future. Another is to look at technologies that are available today, but to focus on bare metal control, cutting out the layers of software that commercial clouds use to manage their resources. It is important to consider not only compute resources, but network, storage, and others as well. A third is to consider performance reproducibility, which has different needs than the performance SLAs offered by commercial providers; generally, performance reproducibility requires little or no sharing of resources. Whatever the specific strategy is, the goal of a cloud computing testbed should not be to duplicate the features of a commercial cloud: it should be to provide those features that are not found there.

Position Paper On Desired Capabilities Of Next Generation Shared Experimental Cloud Computing Research Infrastructures

Joe Mambretti

With its national and international research partners, the International Center for Advanced Internet Research (iCAIR) at Northwestern University designs, develops, implements, and operates large scale, including world-wide, computer science testbeds. Generally, with its research partners, iCAIR operates between 25 and 30 national, international, and local testbeds. The majority of these have been developed as network research testbeds. However, a number of them are based on compute architecture, including the NSFCloud Chameleon, several computational science clouds, and computational science Grid facilities.
Emerging capabilities that will be required by experimenters include those which are needed to support a wide range of key research topics. One is a need for researchers to access a broad diversity of hardware and software architecture, including multiple types of processors and co-processing capabilities, CPUs, GPUs, FPGAs, hybrids, etc. Other importance cloud research topics include enhanced process virtualization, techniques for deeper levels of software defined infrastructure, more flexible and programmable software stacks, orchestrators, tools for managing experimental workflows, and tools for recording, analyzing, diagnosing, and reproducing experiments.

Another emerging need is for AI/ML/DP capabilities at all levels, embedded into almost all infrastructure and processes and as inherent part of the foundation architecture. Also, increasingly, cloud environments must be able to scale to address effectively exponentially increasing amount of data, including at hyper scale. Another need is for federation among cloud testbeds as well as with other types of computer science testbeds, e.g, network testbeds, so that multiple types of resources can more easily be integrated with cloud testbeds.
With its research partners, iCAIR has a special interest in several key topics related to a) using clouds to support large scale data intensive science, much beyond the current capabilities of commercial clouds b) more closely integrating clouds with specialized storage systems for data intensive applications, c) highly virtualized cloud tenant networks, d) proving more programmable L2 and L1 capabilities in cloud environments, and using P4-based telemetry techniques to obtain high-fidelity visibility and analytics in cloud networks.

Other special topics include integration of cloud environments with large flow networks. With its partners in the Chameleon project, iCAIR implemented a large data flow appliance (100 Gbps) for experimenters. On experimental network testbeds, iCAIR has been developing 400 Gbps and Tbps networking capabilities, including for DCIs. Providing these capabilities to cloud researchers would be particularly useful to enable them to prepare for managing extremely large scale data flows to and from clouds.

Wireless Ecosystem meets the Cloud

Abhimanyu Gosain

Our previous experience operating distributed cloud platform(s) such as Global Environment for Network Innovations (GENI), now managing city-scale wireless research testbeds; COSMOS and POWDER-RENEW as part of the Platforms for Advanced Wireless Research (PAWR) program and a DARPA RF emulator Colosseum operating in a cloud environment has allowed us to observe the democratization of cloud access for various science(s), industry(s) and federal agencies. Increasingly, we are realizing that wireless research is complex and intertwined with multiple disciplines; radio hardware, software, systems, networking and cloud. We ask ourselves how we combine all of the above with data and applications. Our perspective for this workshop is to share experiences of the wireless ecosystem, its needs from the cloud community, and share lessons learned from providing cloud access to multiple stakeholders. This ecosystem is nascent in its adoption of cloud technologies and bringing wireless control and user plane data into the cloud is still not mainstream. The adaptability of cloud to scale up and down as service needs change across multiple trust domains; edge, core, IXP or on-premises private cloud(s), is an active area of research.
5G, NFV and software defined infrastructure have been the driving forces behind cloud adoption in the wireless ecosystem for the last few years. However, the full benefits of intelligent decision-making and information sharing in a cloud environment for the wireless ecosystem will be driven by AI frameworks. This is dependent upon cloud technologies and hardware variants‚Äô (GPU, TPU and P-4) ability to provide visible, accessible, understandable, and trusted data in a timely manner.

Federation/Sustainability of Facilities

Our managed facilities combined with other national federally funded testbeds have a sustainability metric for long term continued operations. This means charging users for access. Along with the above, federation to allow testbed user(s) access across various domains to the portfolio of federally funded research infrastructure is crucial. These requirements demand a uniform Identity, Credential, and Access Management (ICAM) policy among the various facilities. It is early days for the facilities including cloud focused ones to onboard paying customers but we would like to share our approach on this topic. ICAM allows for researcher’s digital identities to be exchanged securely and a consistent approach for maintenance of associated attributes, credential issuance for person/institution entities, authentication using the credentials, and making access management control decisions based on authenticated identities and associated attributes. This uniformity of access allows multi-disciplinary users and even the facilities themselves to leverage each other’s resources.

ECAS: Exploring Clouds for Acceleration of Science

Ana Hunsinger and Jamie Sunderland

In November 2018, NSF awarded Internet2 with a cooperative agreement to implement ECAS (https://www.internet2.edu/vision-initiatives/initiatives/exploring-clouds-acceleration-science/) The major goals of the E-CAS project are to coordinate, facilitate, and oversee a project on behalf of the National Science Foundation (NSF) to accelerate scientific discoveries by leveraging advancements and novel technologies in commercial cloud. The project includes partnerships between NSF, Internet2, and Amazon Web Services and Google Cloud Platform as representative commercial cloud providers. The E-CAS project is demonstrating the effectiveness of current commercial cloud platforms in supporting a range of applications critical to growing academic/research computing and computational science communities and illustrating the viability of commercial clouds as an option for leading-edge research computing spanning a broad spectrum of scientific disciplines. By doing so, the project is helping the broader research community understand how the simulation and applications workflows currently using NSF's High-Performance Computing (HPC) and High Throughput Computing (HTC) resources, as well as hybrid campus cloud platforms could benefit from the scale and newer innovations of commercial cloud platforms. To address this, the project is exploring how scientific workflows can innovatively leverage advancements in real-time analytics; accelerated processing such as graphical processing units (GPUs) and field-programmable gate arrays (FPGAs); and automation in deployment, scaling, and management of containerized applications, in order to provide digital research platforms to a wider range of scientific endeavors.

The project has focused on acceleration of science (to achieve the best time-to-solution for scientific application/workflows using cloud) and innovation (to explore innovative use of heterogeneous hardware resources to support and extend application workflows. This is occurring within two phases. Phase I (April 2019-April 2020) has included the review and selection of six Phase I awards, each with a one-year period to develop their projects. The project,'s second Phase will begin in June 2020 with selection of two one-year awards from proposals received from Phase I awardees.

Though ECAS is still in its early stage, there are a number of lessons learned that we believe would be important to share with the broader community along with an update on the current projects.

In addition to ECAS and if appropriate, Internet2 is willing to share a short summary on other current activities in support of research and cloud computing.

Next Generation Cloud Infrastructure and Platform for Mission Critical Enterprise Applications

Carlo Bertolli

Enterprise workloads support critical business processes for companies operating across sectors such as financial services, health services, transportation, and telecommunications. Current market research suggests that only 20%[1] of enterprise workloads have migrated to the cloud. Examples of migrated workloads include experiments, for instance, data analytics using innovative AI techniques deployed to a public cloud; or user-facing apps, where a front-end mobile app is deployed to the cloud and loosely coupled with on-prem traditional IT servers. The missing 80% are what we call enterprise mission critical applications, i.e. applications with strict reliability, performance, and security constraints. Additionally, studies show that 50% of enterprise applications, which are still to migrate to the cloud, operate in regulated industries (e.g. financial services). Top concerns for the move of enterprise mission critical applications to the cloud are security and compliance to regulations (e.g. GDPR, CCPA). Much innovation is still needed to address this.

First, new mechanisms for security are needed at the infrastructure level. Hardware and operating system co-design is necessary to provide ubiquitous, impossible to circumvent, and fully auditable data confidentiality and trust. We will investigate AI-driven vulnerability detection to reduce day-0 vulnerabilities in our platform. We will also strengthen Linux isolation capabilities with light-weight kernel built-in mechanisms, based on address space isolation, to make containers isolation equal to or better than that of VMs. Our platform will leverage hardware security modules to achieve whole stack and application integrity measurement and enforcement. We will leverage these capabilities to strengthen application identity management in cloud -- a necessary capability to enable secure application interconnectivity.

Second, security threats are introduced by software vulnerabilities arising from misconfigurations related to integration of the different application components. For instance, several cross configuration issues and human errors were causes of recent widely advertised breaches [2, 3]. To overcome this, we will introduce an integration focused programming model that enables high-level specification of multi-component solutions and generation of low-level resource and platform configurations. With this abstraction we will be able to automatically generate configuration code, reducing the chance of human errors and introducing the possibility of extended automatic checks based on policies in the CI/CD pipeline.

Third, a transformation of the application compliance processes is needed. Application compliance management encompassing vulnerability-, integrity management, data access protection, and policy enforcement, is an essential concern for enterprises. Our goal is to provide the enterprise compliance persona a platform with enterprise-wide consistent visibility to the compliance state, and risk-based assessment of the compliance posture. The platform will continuously collect the required evidence and assesses and quantifies the findings to advise on the risk and prioritize automatable actions. Audits will no longer handled manually or in an ad hoc manner. Service providers will manage compliance based on the actual risk allowing them to prioritize remediation actions. The platform leverages control points optimized for cloud native environments, where validation is shifted left to pre-deployment phases, and workloads are deployed into a trusted secure environment that is guaranteed to protect against threats in multiple levels.

[1] Think 2019: IBM CEO Ginni Rometty on Chapter 2 of Digital Transformation. https://newsroom.ibm.com/think-spotlight?item=30994, (Aug 2019).

[2] How the Equifax hack happened, and what still needs to be done. https://www.cnet.com/news/equifaxs-hack-one-year-later-a-look-back-at-how-it-happened-and-whats-changed/, Aug 2019).

[3] Capital One data breach involves 100 million credit card applications. https://www.cnet.com/news/capital-one-data-breach-involves-100-million-credit-card-applications/, (Aug 2019).

CloudBank

Rob Fatland

Under NSF support and direction the University of California San Diego, the University of Washington and the University of California, Berkeley will develop and operate CloudBank, an entity that will help the computer science community adopt the public cloud as a research and education computing platform. CloudBank will combine managed services with education and training resources to demonstrate cloud viability in research computing. CloudBank services will address a spectrum of cloud user needs including front line user support, cloud solution consulting, training, and assistance in preparing proposals that include cloud resources. CloudBank will reduce cloud adoption pain points such as managing cost, translating computing environments to cloud platforms and learning cloud-based technologies that accelerate and expand research. It will be complemented by a cloud usage monitoring system that gives NSF-funded researchers the ability to easily grant permissions to research group members and students, set spending limits, and recover unused cloud credits. All of these tasks can be viewed as the baseline of cloud infrastructure. I will touch briefly on both why this baseline is inadequate to realizing the potential of the public cloud in research and how complementary open source projects like pangeo, cloudknot, and beaker are rapidly filling out the necessary upper tiers of the cloud cyberinfrastructure technology stack.

Unidata in the Cloud

Mohan Ramamurthy

This position paper for the Next Generation Cloud Research Infrastructure (NGCRI) is submitted from the perspectives of Unidata, a geoscience cyberinfrastructure facility for the Earth system science community. Earth System Science involves studying the processes and interactions among the atmosphere, hydrosphere, cryosphere, biosphere, and geosphere from a global to local point-of-view, and across the time scales in which these spheres interact. Unidata aims to transform the conduct of research and education in the Earth system sciences by providing well integrated data services and tools that address the entire scientific data lifecycle, from locating and retrieving data, through the process of analyzing and visualizing data either locally or remotely, to curating and sharing the results. In doing so, Unidata strives to lower barriers to accessing data and tools, reduce data friction, and thereby shrink time to science.

The Unidata Program Center is sponsored primarily by the National Science Foundation. Several hundred institutions worldwide participate in the Unidata data sharing network and many more institutions use Unidata tools and technologies in education, research, and operations.

As the enabler of a broad community, Unidata

- Acquires and distributes data to facilitate Earth System Science and education
- Develops software for accessing, managing, analyzing, visualizing, and effectively using those data
- Provides comprehensive support to users, including training.

Unidata has been developing and deploying data infrastructure and data-proximate scientific workflows and analysis tools using cloud computing technologies for accessing, analyzing, and visualizing geoscience data. In addition to fostering the adoption of technologies like pre-configured virtual machines through Docker containers and Jupyter notebooks, other computational and analytic methods are enabled via “Software as a Service” and “Data as a Service” techniques with the deployment of Unidata tools and services in the cloud. The collective impact of these services is to enable scientists to use the Unidata Science Gateway capabilities to not only conduct their research but also share and collaborate with other researchers and advance the intertwined goals of Reproducibility of Science and Open Science, and in the process, truly enabling “Science as a Service”.

Unidata has implemented the aforementioned services on the Unidata Science Gateway which is hosted on the Jetstream cloud, a cloud-computing facility that is funded by the U. S. National Science Foundation. The aim is to give geoscientists an ecosystem that includes data, tools, models, workflows, and workspaces for collaboration and sharing of resources. Unidata’s Science Gateway aims to bring sophisticated cloud-computing solutions to universities without requiring them to develop extensive in-house expertise. It lets educators take advantage of cloud platforms that are pre-configured with all of the tools and data they need.

The vision for NGCRI must include playing a fundamental role in the education, training, and workforce development of the next generation of geoscientists in modern data-intensive modes of scientific investigation and human capacity building. At present, significant barriers to entry exist in the use of cloud infrastructure for education.

Numerous studies and reports have raised alarm that women, racial and ethnic minorities, and persons with disabilities are underrepresented in the STEM disciplines, and geoscience in particular. Broadening participation in the use of research cloud infrastructure should be an explicitly stated goal and an integral aspect of the next generation cloud infrastructure. Through research and practice, that infrastructure must strive to ensure that outreach and support systems are in place to build the skills, including people of all races and ethnicities and underrepresented minorities, and those with disabilities. Specifically, proactive outreach to and engagement of individuals from Minority Serving Institutions, Historically Black Colleges and University, Hispanic Serving Institutions, and Native American Colleges will be needed to make progress and achieve success in this difficult area.

A few other outstanding issues from the Unidata perspective:
a) A cost-effective business model for scientific cloud computing resources
b) A robust, reliable and efficient cloud environment for the academic community, particularly for quasi-operational services like Unidata's services that ingest and provide over a terabyte of real-time data.
c) Persistent services vs. batch computing
d) Integration of HPC and cloud resources, an area that is ripe for fruitful advances in the future.

Choosing the Right Programmable NICs for Future Shared Cloud Infrastructures

Brent Stephens

Networks are beginning to support complex offloads and in-network computation to accelerate applications, reduce the load on general purpose CPUs, and provide new complex and stateful network functions. Programmable NICs have emerged as a crucial enabling technology for in-network computing. Thus, next generation shared experimental cloud infrastructures should be provisioned with programmable NICs so as to enable researchers to develop and experiment with new offloads and in-networking compute use cases. Unfortunately, programmable NICs are also difficult to program and integrate into existing applications. This is a significant barrier to entry, even for the research community. As such, there is also a need for the creation of new open-source toolchains and frameworks to enable researchers to most effectively create their new systems and fully utilize available technologies.

Deciding which types of programmable NICs to include as part of future shared cloud infrastructure is difficult. There are many different programmable NICs, each with their own benefits and trade-offs. However, two types of NICs stand out in particular as important for inclusion as part of future shared cloud infrastructure: Manycore NICs, and FPGA NICs. Manycore NICs use many embedded CPU cores located on the NIC to provide offloads and in-network computing, while FPGA NICs use an FPGA substrate to provide offloads and in-network computing. The primary benefits of Manycore NICs are that they are easy program and manage as they can run standard OSes like Linux. However, the major drawback associated with Manycore NICs is that they incur high latency and have limited per-core throughput. In contrast, FPGA NICs can scale to 100Gbps line-rates and beyond with low latency. However, FPGA NICs are difficult to program and difficult to share between competing applications. This is because these NICs are programmed in low-level languages like Verilog and require expertise in hardware design to extract peak performance from the NIC.

Any shared cloud infrastructure should be able to support a large number of diverse applications and be usable by researchers with different expertise. Because of the above trade-offs associated with Manycore NICs and FPGA NICs, a next generation shared experimental cloud will likely need to provide access to both Manycore and FPGA NICs to meet the requirements of different researchers and different applications. This is in part because choosing the wrong platform or offload design can harm performance. For example, if only Manycore NICs are available, then a researcher may incorrectly conclude that it is not possible to accelerate a latency-sensitive application with a programmable NIC when in fact an FPGA NIC could. This is particularly problematic in the context of creating a shared cloud infrastructure. If only the wrong type of programmable NIC is available, researchers may incorrectly conclude that it is not possible to accelerate their application with programmable NICs. On the other hand, if only FPGA NICs are available, then the shared infrastructure would not be usable by many researchers because of the difficulties associated with hardware design. Further, even for the researchers who are experts in hardware design, it still may have been the case that using a Manycore NIC instead would have reduced development time and effort.

However, in the future, it may be possible to overcome the need for multiple different programmable NIC platforms through a community software infrastructure effort on improving FPGA NIC support. Most, if not all, of the problems with FPGA NICs can be addressed with better software toolchains and frameworks. For example, with a new compiler and toolchain, it could be possible to program FPGA NICs in high level languages. Further, because of their programmability, it is possible to program an FPGA NIC to behave as a Manycore NICs, so there is little lost in terms of functionality if only FPGA NICs and Multicore NICs were available.

Unfortunately, so far, there has yet to be any coordinated effort on the front of developing such open-source toolchains and frameworks. For example, even though the Catapult FPGA platform from Microsoft is internally being used as a programmable NIC, it has yet to be successfully used for programmable NIC research because there is no open-source software for using it as a NIC. To ensure the growth that development of in-networking, it is necessary for the community to recognize and address this problem.

National and Local Impacts of Programmable Cyberinfrastructure using Jetstream

David Hancock, Craig Stewart, Jeremy Fischer

Jetstream is the first production cloud funded by the United States' National Science Foundation (NSF) for conducting general-purpose science and engineering research as well as an easy-to-use platform for education activities. Unlike many high-performance computing systems, Jetstream uses the interactive Atmosphere graphical user interface developed as part of the CyVerse project and focuses on interactive use. This interface provides for a lower barrier of entry for use by educators, students, practicing scientists, and engineers and has been a key tool in democratizing cloud access and educating the future workforce. Over 1,000 students have access under current allocations and thousands more through the science gateways that the system supports.

Jetstream is built upon OpenStack and enables programmable cyberinfrastructure with API access to allow orchestrated and dynamic workloads using Kubernetes and other tools. A significant driver of the advanced interfaces are science gateways and data repositories that leverage persistent hosting, API access, and a flexible storage environment. A key part of Jetstream's mission is to extend the reach of the NSF's eXtreme Digital (XD) program to a community of users who have not previously utilized NSF XD program resources, including those communities and institutions that traditionally lack significant cyberinfrastructure resources. Jetstream has been fulfilling that mission to date; 86% of users in the first project year had never executed a job on another XSEDE (eXtreme Science and Engineering Discovery Environment) system and that trend continues into the current project year, year four of five. This has been accomplished through a strong network of partners such as XSEDE Campus Champions, the Science Gateways Community Institute, and holding events in the majority of states in the US (39 of 50 to date). The Jetstream team has also fully embraced the mission of directly educating the future workforce by hosting undergraduate students as part of a summer REU (Research Experiences for Undergraduates) program each of the last 3 years with the majority of support coming via NSF funding supplements. Diversity has been a priority for the Jetstream team with a majority of these REU students being female or minority students, often from smaller institutions.

The team has also embraced the mission to assist others with cloud operations and instantiation at their own institutions through participation in the OpenStack Scientific working group, international collaborations, and helping provision resources through efforts of the XSEDE Cyberinfrastructure Resource Integration group. This work has benefited not only the national community but Indiana University's (IU) local CI environment by continuing to push the boundary of cloud and HPC operations with key technologies that will be featured in IU's latest AI-driven supercomputer, Big Red 200 (a Cray Shasta system leveraging AMD CPUs and NVIDIA GPUs). While not the traditional focus of cloud computing the overlap continues to grow as will be discussed at an upcoming SC19 workshop, SuperCompCloud: Workshop on Interoperability of Supercomputing and Cloud Technologies (chaired by Sadaf Alam from the Swiss National Supercomputing Centre and David Hancock from IU).

Designing and Deploying High-Performance and Scalable MPI Library for Amazon-AWS and Microsoft-Azure Cloud: Opportunities, Challenges and Experiences

Dhabaleswar K (DK) Panda

Cloud providers like Amazon-AWS and Microsoft-Azure are aiming to provide solutions for High-Performance Computing (HPC) users and applications. Microsoft-Azure has recently moved to bare-metal environment with InfiniBand networking technology. Amazon-AWS has moved to using a new network adapter, Elastic Fabric Adapter (EFA), for its HPC instances. These new networks and platforms provide a new set of opportunities and challenges to design high-performance MPI libraries for these emerging cloud platforms.

The MVAPICH MPI library [1] has been a significant component in the HPC domain for the last 18 years. Recently, the MVAPICH team members have worked closely with Amazon-AWS and Microsoft-Azure to design and deploy high-performance and scalable MPI libraries in these cloud presentation.
On the Amazon-AWS side, a new design [2] has been incorporated into the MVAPICH MPI library to work with the new Scalable Reliable Datagram (SRD) transport protocol available with the EFA adapter. This new design performs better than the existing design using the Unreliable Datagram (UD) and also better compared to other vendor provided MPI libraries. The design has been made publicly available [3] recently for the HPC community.

On the Microsoft-Azure side, the existing MVAPICH2 design has been optimized to the bare-metal configuration of the Azure HB and HC instances with virtualization. Multiple optimizations related to point-to-point and collective communication have been carried out for this platform. The design, with support for “one-click” easy deployment done in conjunction with Microsoft, has also been made available [4] recently for the HPC community.

This presentation will provide details on these designs, their performance (micro-benchmarks and applications) and comparison with other MPI libraries. A set of results comparing the performance numbers on cloud environments with dedicated (non-virtualized) clusters will be presented.

[1] MVAPICH: MPI Library for InfiniBand, Omni-Path, Ethernet/iWARP and RoCE, http://mvapich.cse.ohio-state.edu/.

[2] S. Chakraborty, S. Xu, H. Subramoni and D. K. Panda, Designing Scalable and High-Performance MPI Libraries on Amazon Elastic Fabric Adapter, Hot Interconnects 26, August 2019.

[3] Release of MVAPICH2-X-AWS 2.3 for Amazon AWS Cloud, http://mailman.cse.ohio-state.edu/pipermail/mvapich/2019-August/000160.html.

[4] Release of MVAPICH2-Azure 2.3.2 for Microsoft Azure Cloud, http://mailman.cse.ohio-state.edu/pipermail/mvapich/2019-August/000161.htm.

Humans in the Loop - the critical importance of people in the use of cyberinfrastructure in the cloud

Brian D. Voss

This session explores the important role humans in support roles have in making cloud computing useful in research settings. Cloud computing is clearly a type of cyberinfrastructure, which is defined as comprising computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and enable breakthroughs not otherwise possible. This session will focus on the ""and people"" part of cyberinfrastructure, and in particular on the role of people in supporting the use of commercial cloud resources in research.

The overall goal of this session is to share information about best practices, successes, and challenges in supporting research use of commercial clouds. To accomplish this, we will:

- Provide concrete examples of the effective use of cloud computing, comparing and contrasting use of cloud resources with traditional campus and national infrastructure alternatives.
- Spur discussion for a free exchange of ideas, challenges, and best-practices in supporting the use of commercial cloud computing in advancing research across many disciplines.
- Identify interested parties for further exploration of ways to improve both the grasp of the role of humanware in support of cloud use, and create communities of “humanware providers.”

The session will detail the observations and outcomes of an effort led by the Pervasive Technology Institute at Indiana University Humanware’s Advancing Research in the Cloud (or H-ARC). https://humanware.iu.edu. Microsoft has funded H-ARC, but the PTI is agnostic toward cloud vendor offerings and seeks to also detail the broader return on investment (ROI) of all cloud computing alternatives to traditional premise-based and national cyberinfrastructure paradigms.

H-ARC completed its first phase, involving eight (8) separate projects at research-intensive universities in the US. We will share the participants' reflections regarding the role and impact of humanware in cloud computing use, insights as to the ROI of cloud versus premise/national CI alternatives, and also feedback and perceptions of the value -- and shortcomings -- of cloud vendors approach and support of the use of their products in research. As well, we will share plans and progress for expanding the impact more broadly across the US, North America, and linking with global efforts in Europe and the Asia-Pacific regions.

Toward compliance-driven CI to democratize access to research on sensitive data

Ron Hutchins, University of Virginia

The need to secure research data today is complex, and is much more than simply imposing security on the data. Instead, research data protection means adhering to terms and condition specified in the Data Usage Agreement (DUA), which is part of the overall project contract or terms and conditions. DUAs are negotiated between the research institution (on behalf of the researcher) and the data owner to specify the kind of data protection required. Depending on the type of data, DUAs may invoke existing compliance systems prescribed by law, such as HIPAA, FISMA, or specify custom protocols required by the data owners (e.g., terms in a Non-Disclosure Agreement). Data protection systems are often highly complicated. For example, NIST 800-171 (mandated for Controlled Unclassified Information) specifies 110 individual controls. These controls specify compliance well beyond technical security, including topics such as trainings and best practices. Therefore, compliance to the DUA is an institutional process involving multiple participating entities, such as sponsored programs, legal, and information security.

In order to meet compliance for protecting sensitive research data, research computing infrastructure must be able to coordinate tightly with research needs and relevant functional unit across the institution. As the result, compliance-capable computing infrastructure are often purpose-built (i.e., supporting a single protection system), making them an extremely expensive investment for universities. For example, the University of Virginia has only one protected high performance cluster, which is HIPAA-compliant to support its health research programs. With new data protection systems continuing to come online, the burden of building compliant research computing infrastructure is quickly becoming prohibitively expensive. This challenge is exacerbated at smaller, minority-serving, or teaching-oriented institutions, excluding their researchers from many important research opportunities.

Toward providing compliance-capable cyberinfrastructure and democratize access to research on sensitive data, the University of Virginia is leading a consortium of eleven public universities in Virginia (UVA) to develop an innovative shared research computing cyberinstrument. ACCORD is a new model of CI that leverages virtualization and the latest tools in securing infrastructure to deploy custom compute resources meeting researchers's data protection needs. ACCORD provides customizable secured research environments that can be integrated into each institution's workflows and security models to support the institution's compliance case. In this presentation, we describe UVA's effort to design and develop ACCORD. Our teams leverage knowledge and best practices learned from the Ivy (HIPAA-compliant) infrastructure, as well as our work to extend Ivy's capability to be FISMA-compliant.

Topics of the presentation include:
-Mapping individual controls to mechanism and institutional policies;
-Design and implementation of secure workflows;
-Issues to consider for scaling across multiple institutions (federated access);
-Securing end points, including assuring levels of authority; and
-Compliance best practices.
We will also describe the plan for ACCORD, with is planned to begin development in the Fall of 2019.

In addition to driving exciting scientific discoveries with strong societal impact, the Virginia ACCORD program prioritizes research training and education in data security and privacy for students and researchers across the State. We are especially excited to involve minority-serving (e.g., HBCUs) and non-PhD granting institutions. In this presentation we will also describe our plans for engaging underserved and underrepresented student populations.

They did it and I reproduced it!: Practical Scientific Reproducibility in the Cloud

Kevin Tyle

Cloud computing, hereafter “the cloud”, offers to the community a workspace where multiple users can easily interface with datasets; i.e., they can perform analyses, run diagnostics, form and confirm/reject hypotheses, and create visualizations.

There now exists a rapidly developing/maturing set of tools (such as Kubernetes and Jupyterhub) that can rapidly deploy the software infrastructure necessary to permit a reproducible, data-centric research workflow. The software infrastructure typically consists of free- and open-source packages such as xarray, dask, cartopy, and hvplot. This software ecosystem is currently being well-exploited by the currently-funded Pangeo (https://pangeo.io/), as well as the previously-funded Big Weather Web (BWW; http://bigweatherweb.org/Big_Weather_Web/Home/Home.html) projects. Pangeo and BWW are but two examples of NSF-funded projects that leverage cloud computing resources to realize a broad goal of the scientific community: namely, the need for reproducibility in scientific research.

If reproducibility is going to be practiced in a research setting, then it is incumbent among us to teach it in the educational setting as well. In a typical graduate atmospheric sciences program, such as UAlbany’s, it is standard practice to include refereed journal articles during a typical semester-long course. These articles might be assigned to all; or, perhaps, a student or group of students may be working through a specific set of papers as part of a project. While certainly not all atmospheric science papers follow this model, certainly a significant percentage of publications over the past few decades have involved the use of datasets stored on electronic or magnetic media, and are analyzed and visualized via the use of software. Although exceptions exist, it is rare at present for a reader to be able to reproduce the steps that the original writers took to create their substantive data-centric content. Consider for example a study on rapidly-intensifying winter storms. How did that table of model versus observed mean-sea level pressure get generated? How did the maps which showed the relevant (thermo)dynamical forcing mechanisms get produced? Thinking back to my grad school days, I would have found it so much more edifying were I able to go step-by-step and reproduce these types of results myself!

In this short presentation, I will present a sample (and simple-to-execute) cloud-based workflow, based on my experience in the BWW and Pangeo projects, which reproduces key results of one of a series of three 1990s-era American Meteorological Society journal articles focused on Superstorm 1993, on which I was a co-author. Recognizing that we now have access to superior gridded data sources, such as reanalyses produced by NOAA and the European Center for Medium-Range Weather Forecasts, this will necessarily be an update as well as a reproduction. I will be using resources provided to the BWW project by NSF’s XSEDE program, for which I currently serve as one of UAlbany’s two Campus Champions.

Cloud Computing at NASA's Frontier Development Lab*

Mark Cheung, Andres Munoz-Jaramillo, Paul Wright, Asti Bhatt, Ignacio Lopez-Francos, Atilim Gunes Baydin, Piotr Bilinksi, Daniel Angerhausen & Miho Janvier

Each summer, FDL brings together teams of domain experts and machine learning scientists / engineers to work intensively for eight weeks to tackle some of the biggest challenges in space science, space exploration, and planetary protection. FDL solutions often require the training and deployment of deep neural networks, which are typically carried out on commercially available cloud compute infrastructure contributed by industry partners such as Google Cloud, Intel, IBM and NVIDIA. While FDL teams are co-located during the summer, collaborations persist for many more months, resulting in refereed journal, conference, and workshop publications and/or presentations.

In this talk, the mentors of teams at NASA FDL and FDL Europe** will present case studies of how FDL teams use cloud storage and compute technologies for data preparation, rapid prototyping, and for scaling scientific and machine learning workflows to hundreds and thousands of machines . We also discuss how FDL teams use online tools (e.g., GitLab, Slack, Google Docs, Dropbox Papers) to facilitate effective remote collaboration. The domain areas covered in our case studies include astrobiology, exoplanet detection, space weather, lunar exploration and astronaut health monitoring.

*NASA's Frontier Development Lab (FDL) is a research accelerator supported by NASA, the SETI Institute and industry partners.

**FDL Europe takes place concurrently with NASA FDL and is hosted at the European Space Agency and the University of Oxford.

Solid Clouds - Clouds Quick and Dependable Enough to be a Seamless Part of Reality
Glenn Ricart

An important class of future clouds serving the Smart and Connected Communities of the Future are those whose actions interact seamlessly with what we now know as the real world. Far from being vaporous or foggy, they will need to be solid enough that society can depend upon them.

- If we are to save the energy lost and pollution gained from stopping and starting vehicles at intersections, we'll need real-time communication among vehicles and solid clouds that makes it appear in the real world that the vehicles just magically know when to take their turn in the intersection with minimal changes in speed.
- In an emergency, my phone (or cochlear implant) may consult a solid cloud to give me personalized directions to exit the building, optimizing everyone's exit path without crowding any single exit.
- Solid clouds may help a distributed team share a single virtual workspace whose apparent physical and cyber-support characteristics promote team imagination, discovery, collaboration, and productivity in a natural way.
- Solid clouds could detect dangerous situations in the making and help us avoid injury, from kids leaning too far over railings to slipping on unseen ice.

The unique properties of solid clouds would seem to be: (a) Dependable enough for the situations they support; (b) Quick enough that their responses and actions don't slow down humans (no waiting time) and are responsive enough for the physical actions they support. As an additional benefit, solid clouds may be quick enough to orchestrate actions and behaviors that appear normal to humans, completely in sync with the rest of the real world.

It's probably not reasonable to have separate digital real-time systems for each set of sensing, actions, and behaviors in the long run -- they will need to integrate with each other as well as the rest of the environment. Those that need to be close to the rest of the real world (to avoid speed of light delays) should usefully be in a nearby solid cloud. In that situation, the different cyber-physical models can fluidly interact with each other without significant communication latency. Maximum resources can be applied to emergencies and dangerous situations in a solid cloud in a digital version of humans having tunnel vision when a life-or-death situation occurs. Similarly, human and assistant interactions of the future should be quick enough that a latency version of the Turing test can be passed. These interactions are cyber-human systems.

One might expect a solid cloud to have redundant sensors and an ability to bypass non-functioning (injured) components and react to a wide range of environmental stimuli. Cloud Orchestration in a solid cloud takes on an additional meaning: it must also orchestrate with the rest of the real world.

Solid clouds are likely to be local (near to the stimuli and actions) and hierarchical and distributed, just as the human nervous system has both peripheral nervous system components and a central nervous system. And, people collaborate in both physical spaces (playing tennis or sharing a virtual white board) and non-physical spaces (enjoying the mind's imagery created by reading a good book or clarifying requests made of a digital assistant).

We need considerable research and experimentation to start merging existing work on any one of these topics into quick-enough and dependable-enough solid clouds.

Chameleon: From Testbed to Ecosystem

Kate Keahey

The primary objective of open testbeds is to make the broadest possible set of experiments accessible to the largest possible set of investigators. Chameleon achieves this by implementing a scalable production testbed (testbed-as-a-service), i.e., providing a set of production services capable of creating and managing many temporary `breakable environments', composed of distributed compute nodes, networks, and storage units, used for individual experimentation. Users are given low level access to hardware as well as a wealth of shared digital artifacts, such as images or orchestration templates, that automate environment deployment and configuration so that complex experimental environment can not only be efficiently deployed, but so that this deployment can be reliably and accurately replicated. Deployed over large-scale and diverse hardware resources, such experimental infrastructure is capable of supporting a large number of flexible experimental scenarios.

Although the primary purpose of experimental testbeds is to provide resources to users who would not be able to satisfy their experimental needs otherwise, an important side-effect is that multiple users now have access to the same resources, compatible with the same digital artifacts. This creates conditions which allow users to share experiments and replicate each other's work more easily and creates an opportunity to foster good experimental practices as well as create a sharing ecosystem. In other words, the experimental platform now provide a "common denominator" for the work of many users: it can be used by reviewers to repeat and evaluate experiments in a paper submitted to a conference, or investigators to replicate and more easily extend each other's work. To facilitate such sharing we implemented a number of integration and sharing mechanisms in Chameleon, ranging from the management of versioning, to integration with computational notebooks, through various sharing mechanisms; all with the objective of allowing users to publish and consume each other's work.

An important aspect of ecosystem development is the ability to engage other providers: effectively packaging the "recipe for the testbed" and thus simplifying both its deployment and operation makes joining a testbed widely accessible. Unlike traditional Computer Science experimental systems which have generally been configured by technologies developed in-house, Chameleon adapted OpenStack, a mainstream open source cloud technology, to provide its capabilities. In addition to defining a position in the debate of whether Computer Science testbeds can be supported on clouds, it also provides a range of practical benefits including familiarity for both users and operators as well as interaction with a strong development community. Working from this base, we added customizations needed to implement a viable platform for Computer Science, as well as many tools that make the operation of Chameleon easier and more cost-effective. The resulting system, packaged and released as CHI-in-a-Box has been deployed at Northwestern University which joined Chameleon as an Associate Site.

In this talk, I will argue that engaging user and provider communities in the work of the testbed can make testbeds both more powerful and more productive. In particular, testbeds are more than just experimental platforms but can serve as a "common denominator" that can eliminate much complexity that goes into systematic experimentation, sharing, and reproducibility.

Computing across Clouds, Campuses, and Scientific Facilities

Paul Ruth

Many scientific research communities and institutions are adopting the cloud as part of their platform to support their computing and storage needs. Rapid adoption of scientific computing models that distribute data and computation geographically and across domains is made possible because of the simplicity by which an individual researcher can obtain resources made available by clouds, campuses, and other scientific facilities. These trends have led to the creation of large distributed scientific facilities deployed across cloud, campus, and administrative domains. Example facilities span those that extend campus IPs into a public cloud using direct layer2 connections (i.e. hybrid clouds), up to complex dedicated infrastructure that are deployed across many campuses, computing centers, data repositories, and public clouds on behalf of a specific science community (e.g. ESnet Superfacilities).

As these distributed facilities become more common and begin to use more sophisticated cloud offerings, researchers will need to experiment with complex distributed systems connected to real clouds, campuses, and local scientific facilities. These experiments will need to be orchestrated across independent public and private resources. Many experiments will investigate network protocols and architectures that are not possible without deeply programmable core networks (e.g. P4, compute, and storage in the network). The cloud and networking research communities must develop ways for researchers to efficiently and securely deploy these novel applications and architectures across disparate compute, storage, and network providers.

This philosophy has driven much of our work on ExoGENI, Chameleon, and many other NSF and DOE projects. It has led to the definition of stitchports, which allow provisioning layer2 circuits between ExoGENI slices and external facilities such as campus, computing centers (e.g. NERSC and TACC), other testbeds (e.g. Chameleon), and even public clouds (via Internet2 Cloud Connect). We have used this technique to successfully deploy large experiments spanning many public and research clouds, software defined networks, and scientific compute facilities. However, GENI is essentially an edge cloud, and there remains a need for a future core networking testbed that is designed to enable researchers to easily deploy experiments across the boundaries of cloud, campus, and scientific facilities. It is essential for this future testbed to include a deeply programmable core and opt-in connectivities to existing cloud, campus, and science facility infrastructure.

The Mass Open Cloud (MOC)

Peter Desnoyers, Orran Krieger

The MOC is a cloud being developed by a partnership of academia (Boston University, Harvard University, Northeastern University, Massachusetts Institute of Technology, and the University of Massachusetts), government (Mass Tech Collaborative, USAF), and industry (Red Hat, Intel, Two Sigma, NetApp, Cisco) to support the growing need for non-HPC resources for research and educational institutions. The MOC is intended to provide a shared environment for machine learning applications and platforms that are expected to drive discovery, and be a target for streaming data from a wide variety of emerging applications and IoT devices involved in data-driven discovery.

The existing MOC physical infrastructure includes around 3000 cores of commodity Intel compute, 44 Power9 cores, 46 GPUs, and 1.2PB of storage. Services offered by the MOC include an OpenStack IaaS service, Ceph Volume and Object storage, an OpenShift/Kubernetes PaaS service, and an experimental AI platform (Open Data Hub). The MOC has been used by thousands of students and researchers over the last five years, supporting courses and numerous research projects. It currently has around 400 active users and over 10,000 users of services deployed by its active users, including e.g. Harvard WorldMap; other widely-used services (e.g. Harvard DataVerse) are migrating from AWS to the MOC as well.

The MOC is today housed in the Massachusetts Green High Performance Computing Center (MGHPCC) in Holyoke MA, supporting 750 racks and 15 MW on a single shared floor. The MGHPCC is owned, operated and used by the same consortium of universities involved in the MOC, housing most research computing for those institutions - currently ~300,000 cores and 45PB - with multi-100Gbit/s wide area connectivity.

The MOC has been developing a set of services, Elastic Secure Infrastructure (ESI), to securely move machines between uses, such as HPC clusters and bare-metal systems research usage. It is being upstreamed into existing well-supported open source projects with the assistance of industry partners, and the research IT departments of two institutions have agreed in principle to use ESI to move hardware into and out of the HPC clusters which they manage.

Finally, we are expanding the MOC both within and across institutions: BU and Harvard are creating a production cloud service based on the MOC, and supported by the professional research IT staff from these universities, and we are working to replicate and federate the MOC across other groups of institutions.

Open Cloud Testbed: A Testbed for the Research Community Exploring Next-Generation Cloud Platforms

Michael Zink, Orran Krieger, Peter Desnoyers

The NSF Open Cloud Testbed (OCT) project will build and support a testbed for research and experimentation into new cloud platforms - the underlying software which provides cloud services to applications. Testbeds such as OCT are critical for enabling research into new cloud technologies - research that requires experiments which potentially change the operation of the cloud itself. The new testbed will a) combine proven software technologies from both the CloudLab and the Massachusetts Open Cloud projects, b) combine a research cloud testbed with a production cloud, through OCT’s tight integration with MOC, c) federate with CloudLab, and d) provide programmable hardware (FPGAs) capabilities not present in other facilities available to researchers today. The combination of a testbed and production cloud allows a) larger scale compared to isolated testbeds, b) reproducible experimentation based on realistic user behavior and applications, as well as c) a model for transitioning successful research results to practice. The programmable hardware will be a unique resource enabling investigation into hardware acceleration techniques, research not possible on testbeds available to cloud researchers today, and the community outreach portion of the project aims to identify, attract, and retain these researchers, and to educate them in the use of the facility. The testbed offers a unique sustainability model, by allowing additional compute resources to be dynamically moved from institutional uses into the testbed and back again, providing a path to growth beyond the initial testbed.

This testbed will accelerate innovation in cloud technologies; technologies affecting almost all of computing today. In providing capabilities that today are only available to researchers within a few large commercial cloud providers, it will allow diverse communities to exploit these technologies, democratizing cloud computing research, and allowing increased collaboration between the research and open source communities. The community outreach activities of the project are targeted to researchers who explore complex distributed systems and cloud platforms, spanning a diverse range of backgrounds, institutions, and regions. Software tools will be developed to provide easy and efficient access by these researchers; tutorials, workshops, and webinars will offer training in the use of these tools and the testbed itself. The project will support educating the next generation of researchers in this field, and existing relationships with industrial partners of the MOC will accelerate technology transfer from academic research to practical use.

Data Management in an Untrusted Cloud

Divy Agrawal and Amr El Abbadi

Once upon a time databases were structured, one size fit all and they resided on machines that were trustworthy and even when they failed, they simply crashed. This era has come and gone as eloquently stated by Mike Stonebraker. We now have key-value stores, graph databases, text databases, and a myriad of unstructured data repositories. However, we, as a database community still cling to our 20th century belief that databases always reside on trustworthy, honest servers. This notion has been challenged and abandoned by many other Computer Science communities, most notably the security and the distributed systems communities. The rise of the cloud computing paradigm as well as the rapid popularity of blockchains demand a rethinking of our na√Øve, comfortable beliefs in an ideal benign infrastructure. In the cloud, clients store their sensitive data in remote servers owned and operated by cloud providers. The Security and Crypto Communities have made significant inroads to protect both data and access privacy from malicious untrusted storage providers using encryption and oblivious data stores. The Distributed Systems and the Systems Communities have developed consensus protocols to ensure the fault-tolerant maintenance of data residing on untrusted, malicious infrastructure. However, these solutions face significant scalability and performance challenges when incorporated in large scale data repositories. Novel database design needs to directly address the natural tension between performance, fault-tolerance and trustworthiness.

At UCSB we have been investigating various approaches to address these needs. In the Cloud context, we have been exploring hybrid cloud environments consisting of private and public clouds. Such settings have been widely used by enterprises, however, fault-tolerant protocols have not been adapted for such applications. We are developing a hybrid State Machine Replication protocol to handle both crash and malicious failures in a public/private cloud environment. Traditional replication protocols assume an infrastructure where failures are benign, i.e., crash or fail-stop. We extend this model by considering a mix of failures. In particular, we consider a model that includes a private cloud consisting of non-malicious (either correct or crash) nodes and a public cloud with both malicious and correct nodes. A key insight from this approach is to show how being aware of where different types of failures (crash and malicious) may occur in hybrid cloud environments results in designing more efficient protocols. In particular, some modes of the protocol require less communication phases and message exchanges and hence closer in performance to Paxos, while in contrast to Paxos, which only tolerates crash failures, malicious failures can occur. Other modes are useful for a heavily loaded private cloud or when there is a large network distance between the two clouds while being more efficient than state of the art hybrid protocols.

We are also exploring blockchains, which can be viewed as asset management databases in untrusted infrastructures. In this context, we are developing protocols to achieve atomic commitment among multiple not trusting blockchains. In particular, we consider distributed transactions consisting of sub-transactions where each sub-transaction is executed on some blockchain. Transaction execution should guarantee both atomicity, i.e., either all sub-transactions take place or none of them is executed, and commitment, i.e., all changes caused by a transaction must eventually take place if the transaction is committed.

A Tiered Cloud for Research Infrastructure
Rick McGeer

At the first Cloud Computing Workshop at the University of Illinois, Chicago, one presentation was on the use of AWS for data reduction in physics. The presentation, a case study by one physicist who used AWS frequently, was met with an enthusiastic reception by the audience, who broke into sustained applause when the speaker concluded with “This is great! All you need is a credit card and you can get all your stuff done!,” One attendee was less impressed; he said “You people are crazy! You’ve always had this! Just get an X.25 certificate and you can get stuff done!” The speaker replied, to laughter and cheers: “Gerhard, it’s a lot easier to get a credit card than an X.25 cert”.

There are several lessons from this. First, the non-financial barriers to the use of institutional systems are an incentive to use commercial systems for commodity work; second, those non-financial barriers are typically in place because of the legal and regulatory environment in which public agencies operate, not because of poor technological practice or obdurate administration; third, research infrastructure is special-purpose, generally not well-suited for commodity data reduction and analysis, often abused for that because “it’s free”. The case in this abstract is the NSF should construct a meta-cloud facility that would permit researchers to have access to both commercial cloud resources and research testbeds, so that the right tool can be used for the right job, and workflows involving multiple clouds can be seamlessly designed and executed.

Use cases for the Cloud for Research and Education:
Storage, data analysis, executable publication: this use case is commodity computing, but available globally through the Cloud and without institutional maintenance. A prominent example is Jupyter Notebooks, widely used in the domain and social sciences and data science as a form of executable publication. They have also recently been used for replication and setup of testbed experiments, on Chameleon and CloudLab.

Experiments which can be done on standard infrastructure. These experiments may require scale but neither deep stack access, nor specialized networks, nor extremely wide-area distribution. Good examples include distributed storage systems without tight latency bounds; collaborative backup systems; persistent messaging services; and so on.

Experiments which require host/cloud stack changes. These experiments require access to policy layers not exposed by typical Cloud stacks. These include power monitoring, and cloud allocation experiments for various goals.

Experiments which require distributed infrastructures and very low latencies, but standard network access. Good examples of these are network mapping; content distribution networks; overlay multicast networks; and collaborative visualization and manipulation of big data across the wide area.

Experiments which require physical network modification. Good examples of this are experiments which require high-bandwidth/minimum latency access to a fixed resource.
Each of these classes has an appropriate infrastructure associated with it. For the first two cases, commercial cloud providers offer the most cost-effective, easy-to-use, and efficient solution. The third is the use case for NSF Cloud facilities, and the fourth the use case for PlanetLab, or its modern descendants such as EdgeNet. Experiments which require physical network modification are best run on GENI.

The real cost of each of these classes of facilities (in real terms) are: Standard cloud: vanishing; the Distributed cloud (EdgeNet) is essentially free; programmable cloud infrastructure (CloudLab, Chameleon, GENI) is easily the most expensive of these resources.

The fundamental problem is that expensive and/or inappropriate resources (CloudLab, Chameleon, GENI, EdgeNet) is often used for commodity computing because standard commercial Cloud resources are charged for usage, whereas research infrastructures are “free”. In fact, these resources are scarce and oversubscribed, and should be kept for their intended use cases. A contributing factor is that commodity computing is often used for data reduction, and this is often done in situ because it’s more convenient than moving the data to an appropriate commodity resource.

A solution is to lower the apparent cost of commercial infrastructure by prepurchasing commodity computing in bulk and doling it out to researchers as part of the NSF support, and technological support to have seamless workflows across research and commercial cloud computing. Such a seamless workflow would involve the automated transfer of data to a commercial cloud provider, and a single sign-on to commercial and research infrastructures.

Cloud-Native Data Formats for Big Scientific Data
Ryan Abernathey

Many scientific fields involve thousands of scientists studying a shared, public petabyte-scale dataset. This pattern occurs in climate science, neuroscience, astronomy, genomics, etc. Data-proximate computing is essential for these applications. The commercial cloud is an attractive place to deploy such data environments, due to its scale, elasticity, high bandwidth, and global accessibility. The Pangeo project aims to build reusable open-source software and infrastructure to enable interactive, scalable analysis of large shared datasets. In this talk, I will share experiences developed within Pangeo on challenges related to moving large scientific datasets into the cloud and outline an effective architecture for analysis-ready big scientific data.

The key technical difference between the cloud and legacy environments is the use of object storage as the primary storage medium. Traditionally, scientific applications have assumed that data will be accessed via a POSIX filesystem. Cloud object storage (e.g. AWS S3, Google Cloud Storage, Azure Blob Storage), in contrast, relies on HTTP calls to read and write data. While cloud platforms do offer POSIX-style filesystem services, the object stores are by far the cheapest and most scalable way to store massive datasets. The most effective way to work with object stores is to use “cloud native” data formats. Examples of cloud-native formats are Apache Parquet, Apache ORC, Cloud Optimized Geotiff (COG), and Zarr. These formats share several important properties: (1) they expose metadata efficiently over HTTP, allowing users to know what’s inside a file without actually downloading the whole thing; (2) they support partial reads over HTTP, allowing users to transfer just the desired data and no more (so-called “lazy” access); and (3) they avoid using locks in client libraries, enabling massive-scale distributed processing. Unfortunately, most legacy scientific data analysis libraries and data file formats assume that data will always be accessed over a POSIX filesystem. This leads to inefficiencies and friction in trying to migrate to the cloud.

By leveraging the modular design of our software stack, the Pangeo project was able to reengineer our workflow around a new storage library, Zarr, while keeping the higher level, user-facing components unchanged. Using the open-source software packages Xarray, Dask, and Zarr together in a cloud-computing environment, we have developed environments that enable users to interactively visualize and analyze many TB of shared data in minutes. We have focused on the climate domain, but this design pattern could easily be reused in biomedical imaging, astronomy, or any other domain with large, multidimensional array datasets. Challenges remain in efficiently transforming existing legacy datasets into cloud-native formats, and in figuring out how to pay for cloud storage.

Seven principles for effective scientific big data systems

Joe Hamman, Ryan Abernathy, Niall Robinson

Across many scientific disciplines, from geoscience to biology to astronomy, scientists want to ask questions of very large, complex datasets. However, discoveries in the Big Data era do not come for free; scientists are constantly working to cope with a host of challenges unique to working with large datasets. Cloud computing offers researchers a new opportunity for interactive scientific computing at scale.

In this presentation, we will outline seven principles for effective scientific big data systems that have underpinned the development of the Pangeo Cloud Platform. The Pangeo Project itself is a community effort for open, scalable, and reproducible science. Funded in part through the NSF EarthCube and NASA ACCESS programs, the project has pioneered the deployment of scalable data science platforms for research in the cloud. Although the Pangeo Project is primarily focused on addressing big data challenges in the geosciences, the Pangeo Cloud Platform has been employed by many other domains including finance, astronomy, and biology.

We know that no one data platform will serve the needs of all scientists. Instead we focus on concepts that support the development of powerful, flexible, and efficient data systems. We offer seven architectural principles that we believe are essential in order to create effective, robust, and flexible platforms that make the best use of emerging technology. These principles are:

1) Move data as little as possible
2) Separate concerns and specialize late
3) Scale compute elastically
4) Analyze data lazily
5) Federate data platforms
6) Utilize high-level data brokers
7) Build on open infrastructure

Our presentation pairs these principles with our experience building the Pangeo Cloud Platform to provide a practical vision for the future of scientific computing on the Cloud. We have developed Pangeo in close collaboration with actual domain researchers. This has provided us with important feedback on the challenges scientists encounter when moving their research to the cloud and helped shape the development roadmap going forward. Additionally, it has given us an opportunity to explore the costs of running these systems of commercial cloud systems.

Should NSF Really Host Cloud Infrastructure?
Jeff de La Beaujardiere

This paper takes the possibly controversial position that the National Science Foundation (NSF) may wish to consider, instead of procuring and operating a future research-specific, semi-private cloud, seeking a partnership with commercial cloud vendor(s) to have equivalent computing resources at comparable cost. (I should stress that this is a question of personal interest, not an official position of NCAR.)

Certainly there exist benefits to NSF-funded infrastructure, most notably that it can be made free for researchers and includes science-specific software and support. However, the latter two can be provided regardless of where the actual infrastructure is hosted, and ""free"" does not mean zero cost to the scientific community, only that NSF has covered the cost up front. Furthermore, extending the cloud-oriented analogy of treating a scalable pool of anonymous servers as ""cattle, not pets"" (Bill Baker, Microsoft, 2012), running your own cloud is comparable to owning a ranch. Can we provide all the services we now provide to researchers via NSF-funded cloud without actually owning the servers?

Benefits to using commercial cloud include out-sourcing the infrastructure hosting and operation activities, no fixed limit on the amount of available cores, more storage capacity than is currently provided by NSF cloud infrastructure, the ability to compute directly on other data that may be hosted in the same cloud, and continuously improving compute options and managed services. However, the commercial cloud is not as suitable for specialized High Performance Computing or for continuous workloads.

NSF has already provided some support for use of commercial cloud through the Exploring Clouds for Acceleration of Science (E-CAS) Project, but the funded proposals are specific science activities (see https://www.internet2.edu/news/detail/17078/) rather than an attempt to answer the question, ""Should we even bother to build our own Cloud?""

Questions worthy of consideration include:
* Can we partner with one or more commercial cloud vendors to get equivalent computing resources for the same amount that we would have spent on purchasing and operating our own infrastructure?
* Can we pre-pay to establish a fixed cost or cost ceiling?
* Can we get a good deal on cloud data storage?
* What agreements should be made regarding data egress, including at the end of the project?
* What, if anything, can only be done in an NSF research cloud?
* What can be done better, faster, cheaper, or only in commercial cloud?

Experiments in integrating commercial cloud services with on premise research computing solutions

Chris Hill

Recent emergence of spot/preemptible pricing models from commercial cloud providers has created a competitive compute price landscape for meeting some modern research computing needs. Responding to this, at MIT we have been exploring how to seamlessly integrate multiple commercial cloud provider resources with on premise resources. The goal is to provide cost effective services and elasticity in one meta-environment that can span many different pools of computing capability, native interfaces and charging/economic models. Our brief presentation will describe some details of the approach being explored and some thoughts on research cloud questions the work raises.

Our approach centers on defaulting all access to resources to a single portal that presents a portfolio of systems. A core set of systems under the portal has been configured to provide exclusively container-based execution of workloads. Using technologies from the Checkpoint/Restore In Userspace project (CRIU) we can migrate these workloads live between cloud providers, and between on premise and cloud systems. Virtualized networking allows for TCP connections to stay open during migration, including open network file ports. The approach is proving quite effective in allowing workloads to start in one location and to live migrate as pricing and availability changes. This is allowing us to meet growing demand driven by statistical/AI technology evolution, while keeping capital expenses at a manageable level.

While the approach is successful, opportunities for cloud CS research into deeper system integration in scheduling, checkpointing and migration, and fuller virtualization exist. These could improve system level support for needed intermediate services and boost utility by developing data storage virtualization that is cost effective and can scale. We will outline some ideas in this space where the computer science community could impact directly computational science productivity and economics.

Enabling Next-Generation Image Analytics on Whole Slide Images Using DRAM

Praveen Rao[1], Daniel E. Lopez Barron[1], Deepthi Rao[2], Ossama Tawfik[3], Arun Zachariah[1]
raopr@umkc.edu, dlopezbarron@mail.umkc.edu, raods@health.missouri.edu, ossama.tawfik@gmail.com, az2z7@mail.umkc.edu
[1]*University of Missouri-Kansas City* [2]*University of Missouri-Columbia* [3]*Saint Luke's Hospital of Kansas City*

Whole slide imaging is touted as a disruptive technology in digital pathology and can produce a gigapixel image—a whole slide image (WSI)—in a few minutes by scanning a glass slide with near-optical resolution.[1] There continues to be growing interest in using WSIs for primary diagnosis and consultation[2] as they can greatly improve the workflow of pathologists. From a regulatory perspective, Philips and Leica Biosystems have received FDA clearance to market their digital pathology systems for primary diagnosis in the U.S. This is a major step forward to enable widespread adoption of WSIs for primary diagnosis; however, the technical barriers due to gigabyte-sized WSIs pose a serious challenge for large-scale storage and retrieval of WSIs. In fact, the Memorial Sloan Kettering Cancer Center plans to scan 40,000 slides per month and expects to produce millions of WSIs in the next few years[3]; this will create several petabytes of WSI data. Fast access to small portions of WSIs is highly desirable for next-generation image analytics on histopathology slides using deep learning, which is becoming attractive for automatic diagnosis of cancers. *We claim that I/O is going to become a significant bottleneck for deep learning training when performed on thousands of WSIs*. Thus, developing an effective solution for fast retrieval of tiles is a critical and timely challenge to address.

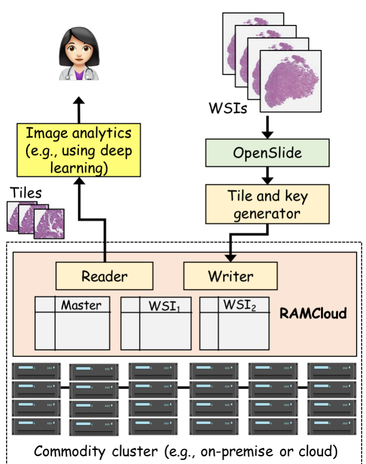

Figure 1: Overall architecture

Our previous work [4] showed that using a space-filling curve for data partitioning and Apache Spark[5] –a popular in-memory cluster computing framework–can enable fast retrieval of (small) tiles in WSIs. Our system's latency to retrieve a title was in seconds when tested on CloudLab[6]. This is because data may need to be fetched from secondary storage at times and move through several software layers. *Given the large amount of DRAM in a commodity cluster, can we reduce the tile access latency if data are always kept in and fetched from DRAM?* We posit this is possible using RAMCloud[7] , a DRAM-based storage system for large datasets with low-latency access. Towards this end, we are developing a new system using RAMCloud's key-value data model, adaptive partitioning of tiles in WSIs using space-filling curves and Dewey labeling scheme for generating keys, and multiread/multiwrite operations of RAMCloud for efficient access. The architecture of our system is shown in Figure 1.

Preliminary evaluation of our system on 100 WSIs[8] using a 24-node cluster on CloudLab showed promising results. (Each node ran Ubuntu Linux 16.04.10 and had a 10-core Intel processor, 480 GB local SSD storage, 250 GB block storage, and 64 GB RAM.) We fetched image tiles of a few MBs in size in millisecond latency. This was in fact 1000 times faster than using our previous work[4] tested on CloudLab using Apache Spark. One of the limitations of CloudLab for scaling our experimentation is the lack of sufficient local storage on cluster nodes. Therefore, we suggest NSF and NIH to invest in experimental testbeds for enabling academic research in digital pathology and precision health as these areas are challenged by massive datasets.

[1] F. Ghaznavi, A. Evans, A. Madabhushi, and M. Feldman. Digital Imaging in Pathology: Whole-Slide Imaging and Beyond. Annual Review of Pathology: Mechanisms of Disease, 8:339–351, 2013.
[2] L. Pantanowitz, J. H. Sinard, W. H. Henricks, L. A. Fatheree, A. B. Carter, L. Contis, B. A. Beckwith, A. J. Evans, A. Lal, and A. V. Parwani. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Archives of Pathology and Laboratory Medicine, 137(12):1710–1722, 2013.
[3] https://www.mskcc.org/departments/pathology/warren-alpert-center-digital- and-computational-pathology
[4] D. L. Barron, D. V. K. Yarlagadda, P. Rao, O. Tawfik, and D. Rao. Scalable storage of whole slide images and fast retrieval of tiles using Apache Spark. In Proc. SPIE 10581, Medical Imaging 2018: Digital Pathology, volume 10581, pages 1–6, 2018.
[5] Apache Spark, http://spark.apache.org
[6] D. Duplyakin, R. Ricci, A. Maricq, G. Wong, J. Duerig, E. Eide, L. Stoller, M. Hibler, D. Johnson, K. Webb, A. Akella, K. Wang, G. Ricart, L. Landweber, C. Elliott, M. Zink, E. Cecchet, S. Kar, and P. Mishra. The Design and Operation of CloudLab. In Proc. of the USENIX ATC, pages 1–14, July 2019.
[7] J.Ousterhout, A.Gopalan, A.Gupta, A.Kejriwal, C.Lee, B.Montazeri, D.Ongaro, S.J.Park, H.Qin, M.Rosenblum, S.Rumble, R. Stutsman, and S. Yang. The RAMCloud Storage System. ACM Transactions on Computer Systems, 33(3):7:1–7:55, 2015.
[8] The Cancer Genome Atlas. https://www.cancer.gov/tcga, 2019.

Leveraging Federation to Enable New Types of Applications

James Griffioen

Although there have been past efforts to federate cloud infrastructure, ""federation"" often (only) means that users have access to the federated cloud resources and can use a common login mechanism (i.e., identity provider) across all the cloud providers. Some efforts have extended the definition of ""federation"" to support common interfaces and APIs for reserving resources (e.g., Cloudlab and Chameleon). Unfortunately, few have gone so far as to agree on and support a common set of APIs for controlling, monitoring, or debugging resources across a variety of different cloud infrastructures.

In recent years, there has also been a growing need among the scientific research community to federate their high performance computing (HPC) computational and storage resources to be able to run larger jobs (spanning multiple institutions), or to ""trade"" access to specialized resources (e.g., one institution's GPU resources for large memory nodes at another institution), or to make effective use of otherwise idle resources. Systems like Open Science Grid, Campus Compute Cooperative, and the Eastern Regional Network have arisen to federate (share) physical resources across institutions, with each approach having its own unique authentication/authorization mechanism, protocols for advertising and requesting resources, and requirements of the physical resources being shared. While some of the HPC federation issues are the same as cloud federation issues, federated HPC systems often have added constraints due to the need to ensure certain network characteristics (e.g., InfiniBand) that are not typically concerns of cloud systems.

It should also be noted that there is a growing interest (and ability) to ""program the network"". The network itself can be viewed as a cloud-like system with computation and storage build into the network that can be reserved and programmed by researchers. As a result, there is an opportunity to federate programmable network systems with federated clouds and/or federated HPC systems. Systems like GENI make this possible on a wide area scale, while technologies like OpenFlow and P4 make this possible on an institutional scale.

In short, recent advances have made it possible to consider federating compute, storage, and network infrastructures together on a local, regional, national, and international scale, revolutionizing the way applications are designed. Applications built from federated systems can now be run everywhere, including the network. For example, future applications might consist of end system client code (like today's applications), but also code to implement the network transport/services used by the application, and even the server code to be run in the cloud or on a federated system with the desired server resources or location.

As a step in this direction, we have been building programmable campus network infrastructure that interconnects our campus cloud resources. In particular, we deployed an OpenFlow-enabled network across large portions of the University of Kentucky campus that can be programmed to dynamically allocate ""VIP Lanes"" -- high performance network paths that are allowed to bypass policy-enforcing bottlenecks in the campus network (e.g., IDS/IDP systems). By ""programming"" the network, we are able to create high-speed channels between the various cloud infrastructures on campus, our HPC systems, and external cloud services (e.g., high-bandwidth channels to Google drive for storage).

We also developed orchestration software that dynamically creates ""personal HPC clusters running OpenHPC"" on one or more of our OpenStack clouds. Users can run a program on their laptop that interacts with our OpenStack infrastructure to create a personal HPC cluster consisting of the number of nodes (VMs), cores per node (vCPUs per VM), and memory per node that the user would like in their personal cluster. The VM nodes are then loaded with the OpenHPC software suite which allows users to submit jobs via SLURM just like a normal HPC cluster. Although the resulting personal HPC cluster lacks Infiniband, the VMs (which can be allocated across multiple OpenStack systems and interconnected by VIP Lanes) are more than capable of running so-called ""single node"" jobs that now dominate our HPC workloads. Moreover, we can make effective use of our large memory cluster where each node has 3TB of memory for big data tasks. Instead of allocating entire physical nodes, our system partitions the memory into multiple VMs right-sized for the current set of jobs (e.g., a personal cluster with 500GB nodes and another with 1TB nodes, both executing simultaneously). Our initial performance tests show that overhead due to virtualization is insignificant.

Intro to Cloud-Ecosystem in Box

Jessie Walker, Brenton Stewart

This talk provides an introduction to the resources/infrastructure hosted within the Cloud-Ecosystem in Box (cloudsforall.org), which is an online collaborative community, initially composed of 10 institutions. Which, is designed to foster collaboration among minority serving institution's faculty/researchers in leveraging cloud computing resources for their research activities.

Leveraging Distributed Cloud Testbeds for Domain Science Research and Experimentation

Anirban Mandal

Computational science today depends on complex, data-intensive applications operating on datasets from a variety of scientific instruments. A major challenge is the integration of data into the scientist's workflow. Recent advances in dynamic, networked cloud resources provide the building blocks to construct reconfigurable, end-to-end infrastructure that can increase scientific productivity. However, applications have not adequately taken advantage of these advanced capabilities. In the context of the DyNamo [4] project funded under the NSF Campus CyberInfrastructure program, we have developed a novel network-centric platform, Mobius [7], which enables high-performance, adaptive data flows and coordinated access to distributed multi-cloud resources (cloud research testbeds like ExoGENI [1], Chameleon [2], XSEDE JetStream [3], etc.), and data repositories for atmospheric scientists.

We have demonstrated the effectiveness of our approach by evaluating time-critical, adaptive weather sensing workflows, which utilize advanced networked infrastructure to ingest live weather data from radars and compute data products on dynamically provisioned resources on hybrid, multi-cloud platforms, which are used for timely response to weather events. The workflows are orchestrated by the Pegasus workflow management system [6] and were chosen because of their diverse resource requirements. We have shown that our approach results in timely processing of CASA [5] weather workflows under different infrastructure configurations and network conditions. We have also shown how workflow task clustering choices affect throughput of an ensemble of workflows with improved turnaround times. Our findings show that using our network-centric platform powered by advanced layer2 networking techniques results in faster, more reliable data throughput, makes multi-cloud resources easier to provision, and the workflows easier to configure for operational use and automation.

We are extending our work such that domain science data flows can be effectively adapted and optimized by leveraging Software-Defined Exchanges (SDX), and the Quality of Service of the end-to-end provisioned infrastructure can be transparently maintained by active monitoring and control. Our current plans also include supporting a wider federation of cloud infrastructure (public clouds like Amazon EC2 and other research clouds like Massachusetts Open Cloud) using Mobius. Using the connected, distributed multi-cloud federation enabled by DyNamo, we are continuing to support (a) a wider range of adaptive weather sensing workflows performing wind computations and hail formation, and (b) ingest of streaming data and on-demand computations for workflows employing data from the Ocean Observatory Initiative (OOI) NSF Large Facility.

References

[1] I. Baldin, J. Chase, Y. Xin, A. Mandal, P. Ruth, C. Castillo, V. Orlikowski, C. Heermann, J. Mills., ExoGENI: A Multi-Domain Infrastructure-as-a-Service Testbed., The GENI Book, pp. 279--315, 2016.

[2] NSF Chameleon Cloud. https://chameleoncloud.org/

[3] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. Scott, N. Wilkins-Diehr, Xsede: Accelerating scientific discovery, Computing in Science & Engineering 16 (05) (2014) 62,Äì74. doi:10.1109/MCSE. 2014.80.

[4] E. Lyons, G. Papadimitriou, C. Wang, K. Thareja, P. Ruth, J. J. Villalobos, I. Rodero, E. Deelman, M. Zink and A. Mandal, Toward a Dynamic Network-centric Distributed Cloud

Platform for Scientific Workflows: A Case Study for Adaptive Weather Sensing, IEEE eScience 2019, San Diego, CA, September 2019.

[5] B. Philips, D. Pepyne, D. Westbrook, E. Bass, J. Brotzge, W. Diaz, K. Kloesel, J. Kurose, D. McLaughlin, H. Rodriguez, and M. Zink., “ntegrating End User Needs into System Design and Operation: The Center for Collaborative Adaptive Sensing of the Atmosphere (CASA)”, In Proceedings of Applied Climatol., American Meteorological Society Annual Meeting, San Antonio, TX, USA, 2007.

[6] E. Deelman, K. Vahi, G. Juve, M. Rynge, S. Callaghan, P. J. Maechling, R. Mayani, W. Chen, R. Ferreira da Silva, M. Livny, and K. Wenger, ,”Pegasus: a workflow management system for science automation”,Future Generation Computer Systems, vol. 46, pp. 17,Äì35, 2015.

[7] Mobius, https://github.com/RENCI-NRIG/Mobius.

Cloud Computing for Improving the Lead-time and Accuracy of High-Impact Weather Forecasts and Warnings in the Western United States

Sen Chiao

Extreme or high-impact local weather occurs daily across a breadth of application areas ‚Äì aviation, surface transportation, harvesting and application of pesticides and herbicides in agriculture, building construction, massive public recreation venues ‚Äì and are characterized by one common factor: they share an inability to apply weather technologies when, where and how needed in order to deal with high-impact weather events. Today‚Äôs numerical weather prediction technology in general, and its use in research, operations and education can be further improved when applied to any particular situation involving high-impact local weather (e.g., atmospheric rivers and wildfires).

It is essential to have an integrated, scalable, web service architecture in which machine learning meteorological analysis tools, high resolution forecast models and data repositories can operate as dynamically adaptive, on-demand, cloud-based systems that will:
(1) configure rapidly and automatically in response to various weather conditions,
(2) respond dynamically to inputs from general users,
(3) enable machine learning processes simultaneously, and
(4) steer artificial intelligence to optimize the data collection for the problem at hand.

Specifically, the Next Generation Cloud Computing will be used on:

- Developing capabilities to allow numerical models and other atmospheric tools to respond dynamically to their own output, to observations, and to user inputs so as to operate as effectively as possible in any given situation;
- Developing appropriate artificial intelligence and machine learning in terms of weather forecasting within supporting the Next Generation Cloud Computing Infrastructure.

The Center for Applied Atmospheric Research and Education (CAARE) at SJSU has been conducting research on numerical modeling, observations, data analysis, machine learning with emphasis on extreme and high-impact weather events all over the world. The overarching goal of CAARE is to advance our understanding of fundamental science in support of the innovative observations for advancing the analysis and prediction of extreme weather events and their linkages to our changing environment.

Converging simulations and learning: Molecular Science Case Studes
Shantenu Jha

The vision of ""Learning Everywhere"" captures the potential and impact of how learning methods and traditional simulation methods can be coupled together. A primary driver of such coupling is the promise that Machine Learning (ML) will give major performance improvements for traditional simulations. Motivated by this potential, the ML around HPC class of integration is of particular significance. We will discuss ``how'' learning methods and simulations are being integrated to enhance effective performance of computations. Studying examples from biomolecular science (see NSF Molecular Sciences Institute: http://molssi.org) we identify several modes --- substitution, assimilation, and control, in which learning methods integrate with simulations and provide representative applications in each mode. We discuss the system software requirements and challenges for cloud platforms that these important classes of applications present.