Abstracts – PEARC20 Workshop Sessions

Sponsor Presentations -- HARC & STRIDES

Pervasive Technology Institute – HARC; Brian D. Voss

Humans in the Loop – The Critical Role People Play in Advancing Research with Cloud Cyberinfrastructure

Ever since technology began to move outside of centralized data centers, people have played an important if sometimes overlooked role in the successful use of technology to advance enterprise and personal productivity functions, teaching and learning, and especially research. Cyberinfrastructure (CI) is often defined as very tangible elements – sensors, hardware, software, networks, visualization, etc.; but it also includes the people who help get the most out of those physical elements. While still not fully grasped by institutions and researchers in today's premise-based CI environments, the rapid advancement of use of clouds – both public and private – brings more complication and complexity to their use in advancing research. The role of people becomes even more critical. The Humans Advancing Research in the Cloud (HARC) project has been examining the broad issues surrounding cloud deployment in support of research, the challenges presented to vendors trying to grasp the unique needs of researchers, as well as researchers struggling to understand cloud services from those vendors not well-constructed for their purpose. A significant element remains the role that skilled cloud research support engineers (CRSEs) play in successfully using cloud alternatives. This talk will review project objectives, accomplishments, efforts underway now, and lessons learned so far.

NIH – STRIDES; Nick Weber

Part 1: NIH STRIDES Initiative

The NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) Initiative allows NIH to explore the use of cloud environments to support the research endeavor by partnering with commercial cloud service providers (CSPs). NIH's STRIDES Initiative facilitates NIH and NIH-funded researchers' access to and use of cloud capabilities and related training and support services from CSPs to help advance biomedical research. As one of many NIH-wide efforts to implement the NIH Strategic Plan for Data Science, STRIDES supports NIH's overall approach to modernizing the biomedical data science ecosystem and accelerating the path to discovery.

Part2: NIH's Managed Cloud Environments

NIH has invested in the development of enterprise-wide cloud platforms and services for internal use across NIH. Key components and capabilities include best practices co-developed and supported by cloud service provider partners, including automation and infrastructure as code; implementation of shared services including integration with NIH login infrastructure well as dedicated, secure network connections from NIH to the cloud; and a baseline of common-sense security controls (e.g., encryption, shared monitoring/logging/alerting, etc.) applied to all cloud accounts or projects within the managed environments. This approach has enabled NIH to centrally offer standardized yet flexible cloud capabilities, reducing duplication of foundational technical components across NIH while allowing for customized implementations by the NIH Institutes and Centers building solutions within the managed environments.

Community Programs Presentations

Internet2 – eCAS; Jamie Sunderland

Community Update -- Internet2 "Exploring Clouds for Acceleration of Science (E-CAS)" project

In 2018, Internet2 in cooperation with the National Science Foundation and representative commercial cloud providers launched the E-CAS project. The project helps researchers understand the potential benefit of larger-scale commercial platforms for simulation and application workflows such as those currently using NSF's High-Performance Computing (HPC), and explores how scientific workflows can innovatively leverage advancements in real-time analytics, artificial intelligence, machine learning, accelerated processing

hardware, automation in deployment and scaling, and management of serverless applications in order to provide digital research platforms to a wider range of science. In the past two years, E-CAS awardees have made significant progress in their understanding of the use of cloud providers; this session will highlight those and point workshop participants toward relevant information on E-CAS progress.

Cloudbank; Shava Smallen, UCSD
Community Update: CloudBank aids the computer science community to access and use public clouds for research and education by delivering a set of managed services to simplify access to public clouds.
CloudBank will serve as an integrated service provider to the NSF computer science research and education community through a set of user-facing and business operations functions. These services include cloud solution consulting and training, assistance in preparing proposals that include cloud resources, and frontline user support. CloudBank will support multiple cloud vendors, and be accessed via intuitive, easy-to-use user portal that gives users a single point of entry to these functions. This talk will give an overview of the CloudBank project, its education and outreach activities, and a brief demo of some of the key features of the CloudBank user portal. Cloudbank is a collaboration of the University of California, San Diego's San Diego Supercomputer Center and Information Technology Services Division, the University of Washington's eScience Institute, and the University of California, Berkeley's Division of Data Science. It is funded by a 5-year NSF Cooperative Agreement (Award #1925001). https://www.cloudbank.org

Papers Received from Call for Participation

IndianaU; James McCombs, Walsh, Noguchi, Kantak
A Hybrid On-premises and Public Cloud Attention Clustering Workflow
Research information technology professionals frequently develop solutions for researchers who need to analyze large data sets. There is a strong cost incentive for utilizing existing on-premises resources, but use of those resources has challenges and risks that can make leveraging public cloud infrastructure a preferable option. Furthermore, it is not always obvious what the challenges and risks will be until an on-premises solution is attempted. We present an interesting case study that arose from a research project in attention clustering in macro and financial economics for which we developed an on-premises solution but realized the solution would not provide the needed performance to ingest and analyze data to keep pace with the rate the data was being generated. As an alternative, we developed a solution that utilized Google Cloud Platform's BigQuery data analysis platform combined with an on-premises analysis front end which provided a much simplified data ingestion and analysis workflow, empowering project members to more easily explore the data, make discoveries and further refine their attention clustering methods. Our experiences demonstrate that under certain conditions, a hybrid solution, which strategically leverages public cloud resources like BigQuery, provides a compelling solution for analysis of large data sets.

Indiana U. – Craig Stewart, et al
Financial returns on federal investment in XSEDE: A half-decade of results
The necessity of computing for research is not in doubt. How much to invest in advanced computing infrastructure, expertise, and other cyberinfrastructure (CI) components is, however, an open question, and is worth exploring at all levels, from institutional to national. The National Science Foundation's funding of the eXtreme Science and Engineering Discovery Environment (XSEDE) is an example of investment in a national-scale provider of CI services supporting national open science research priorities. Thus, XSEDE provides an ideal context for developing and evaluating methods for analyzing return on investments (ROI) in research CI. This paper is the latest of several published over the last five years describing methods of data collection and analysis and the results of analyses of financial returns from NSF funding of XSEDE. It discusses the evolution of the financial ROI model, provides enhancements to methods, and includes new data. The results continue to indicate significant value for the NSF and the nation as a result of funding XSEDE. These results have potential value for consideration of national strategies for cloud computing in the future.

Indiana U.; Marlon Pierce, Marru

Integrating Science Gateways with Secure Cloud Computing Resources

This paper examines scenarios in which science gateways can facilitate access to cloud computing resources to support scientific research using regulated or protected data stored on clouds. Specifically, we discuss the use of science gateways to access Controlled Unclassified Information (CUI), a US regulatory standard that covers a broad range of US federal government-owned or regulated data and which also provides a useful proxy for other types of sensitive data, such as private sector intellectual property. We focus on the impact of CUI requirements on science gateway platforms that can be used to create and manage science gateway instances. Gateway platforms are centrally operated by gateway platform providers, who create and control gateway instances on behalf of gateway providers. Broadly, platforms operate following either a multi-tenant or else a multi-instance pattern. Multi-tenanted science gateway platforms are designed to support multiple science gateways simultaneously, with each gateway as a tenant to a single operational instance of the platform middleware. Multi-instance platforms, on the other hand, provide and manage an entire instance of the science gateway software for each gateway. This paper reviews these two scenarios from the perspective of the Science Gateways Platform as a service (SciGaP), a multi-tenanted gateway platform based on the open-source Apache Airavata software. We examine requirements for providing multi-tenanted platforms for CUI gateways and also the requirements for providing the same software as a multi-instance platform. In both cases, we assume the use of CUI-compatible resources from commercial cloud providers. Both approaches are technically feasible but have tradeoffs that must be considered.

Indiana U.; Rick McMullen

A Strategic Approach to Cloud Services in Academic Research Computing Support

Over the last several years a number of commercial and private cloud services have been made available to host virtual machines and, more recently, containerized application systems. These services range from the commercial giants Amazon, Microsoft and Google, to specialty HPC services like IBM Cloud to smaller niche services like Rescale. That these services need to be considered in the context of a strategic approach to providing campus research computing is not in doubt and there are many examples of successful adoption. What is mainly lacking given the dynamic nature of the cloud market and the diversity of approaches to providing research computing services on college and university campuses is a framework for understanding the issues and tradeoffs in adopting and using cloud services versus equivalent on-premises solutions. This paper lays out some of the human factors and their costs for cloud and in-house solutions.

HARC Program CRSE-2 Participants

ASU; Chris Kurtz

Using industry standard cloud-neutral tools to create a commonly architected, orchestrated, and secure environment that can easily be deployed to any of the common Cloud Providers

ASU's approach to Multi-Hybrid Cloud is to use industry standard cloud-neutral tools including Terraform, Ansible, and Cloud-Init to create a commonly architected, orchestrated, and secure environment that can easily be deployed to any of the common Cloud Providers. Focusing on industry best practices including git and continuous integration/continuous development, cloud environments can be easily deployed, updated, and destroyed as needed. Common to all Public Clouds will be industry standard best practices of utilizing multiple regions and availability zones for critical services, with common security tools across all platforms. VPN connections as well as pre-existing ASU cloud connection products (DirectConnect, ExpressRoute, etc) will provide fast and secure connections to ASU's Research Computing HPC Environment. To begin developing our environment, ASU has chosen to focus on Microsoft Azure. Azure's approach to High Performance Computing is more mature than AWS and GCP, offering HPC-focused products such as CycleCloud and HPC Cache. HPC Cache is of particular interest to provide cost-efficient file syncing in Azure.

GSU; Robert Podschwadt

GSU system architecture for cloud-based privacy preserving machine learning and its challenges and opportunities

Implementing privacy preserving machine learning using homomorphic encryption is computationally expensive. Therefore, running the computations in the cloud is an appealing option. However, it comes with a number of interesting challenges. Most out of the box cloud machine learning solutions are not usable due to the constraints introduced by the privacy preservation. Our system performs training on plain data and inference on encrypted data. Training and inference have different requirements on the underlying hardware. The training phase benefits from GPUs and other hardware acceleration while the inference phase requires lots of CPU cores and large amounts of memories. In this talk, we describe our system architecture for cloud-based privacy preserving machine learning and its challenges and opportunities.

U-Pitt/PSC; Kim Wong

Killing Two Birds with One Stone: Cloud Bursting

The "people" part of cyberinfrastructure is perhaps the most critical component and yet it is the least scalable. People are critical because they provide support that lowers the barrier to adoption of shiny computing/storage systems, advanced instrument and repositories, high-performance networks, and new software. The experts are not scalable because expertise is built from experience and experience takes time. With the emergence of cloud computing, there is concern among my colleagues that we are expected to support users both on our on-premise cluster and in the cloud. Essentially, the other components of cyberinfrastructure are scaling up without a concomitant increase in the number of user support personnel. If only there is a way to make the user experience uniform while on-premise and in the cloud. This talk describes our strategy for leveraging the bursting capability of SLURM to offload computation to commercial clouds and hence, re-use exiting user support channels for the cloud.

Notre Dame; Matt VanderWerf

Case Studies on Migrating Cloud Research Platforms

In this presentation we document three initiatives we have recently undertaken in which software defined infrastructure helped facilitate the movement of projects within and between research clouds. In our first example, we demonstrate the benefits of creating a platform for cloud templates with core-shared services that allows for quick reuse in a variety of compliant environments. Next, we discuss how a cloud implementation for one research partner was taken from proof-of-concept and then formed the basis of a project for a separate partner. The third example involved replicating a time-sensitive project between institutions and how available cloud tools allowed us to expedite the process. In addition, we discuss the importance of the human factor in this endeavor, how we approached each solution, and how collaboration and sound communication are keys to success.