CERIAS

The Center for Education and Research in Information Assurance and Security

User Identity Mapping for Secure Workflows Spanning Cloud and HPC in the Anvil Supercomputer

Sathvika Kotha[1], Erik Gough[2], Rajesh Kalyanam[2]

1. Computer and Information Technology, 2. Purdue IT, Rosen Center for Advanced Computing (RCAC)

Anvil System Description

Anvil is Purdue University's most powerful supercomputer, providing advanced computing capabilities to researchers in many diverse scientific disciplines. Funded by the National Science Foundation (NSF) through a $10 million system acquisition grant, Anvil supports scientific discovery through the NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), providing computing resources to thousands of researchers across the United States.

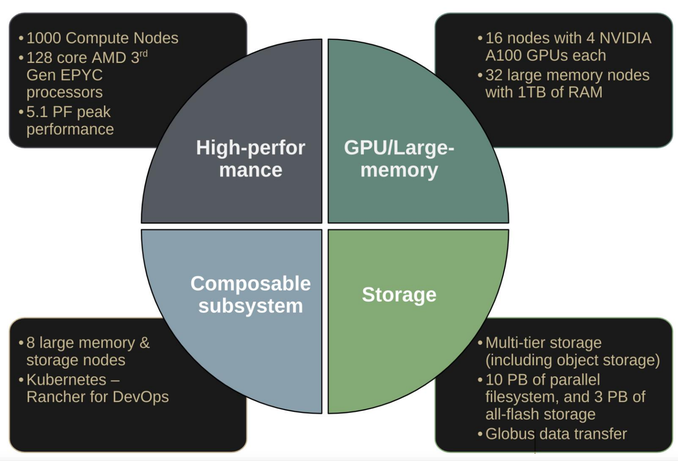
HPC Batch System:

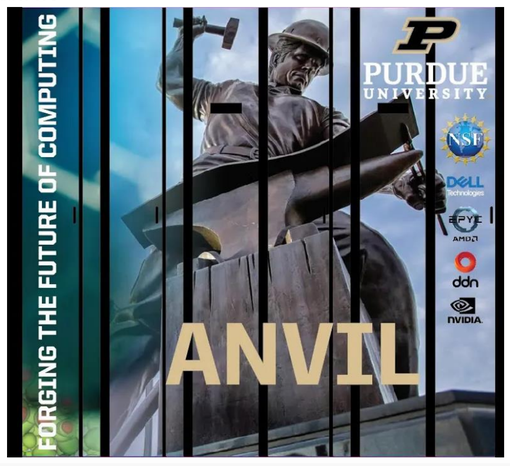
Batch systems allow users to submit jobs requesting required resources (CPUs, RAM, GPUs). Jobs are queued and run as resources become available.

Anvil GPFS storage is a Multi-tier storage system with 10PB of parallel filesystem and 3 PB of all-flash storage.

Anvil Composable Subsystem:

The Anvil Composable Subsystem is a Kubernetes-based based private cloud that provides a platform for creating composable infrastructure on demand. This platform gives researchers the ability to deploy and manage container-based applications and persistent services to complement HPC workflows. Rancher is used as a control plane for Kubernetes cluster management and container provisioning.





Problem Statement

The Anvil Composable Subsystem is a cloud environment ideal for hosting persistent web services such as science gateways which are increasingly providing low-barrier access to HPC resources. Anvil is unique, providing both a composable cloud and HPC system; presenting the opportunity to support seamless data sharing between the two. However, there are security constraints, access control, and authorization challenges that need to be addressed.

In the Anvil cloud, users are able to deploy container-based applications as any user id (uid), even root. Anvil's storage systems use NFS and authorize users based on uid and UNIX permissions. A solution is required that validates a user's ACCESS identity and runs containers as their assigned Anvil uid so data in the storage system can be securely accessed.

Solution

- One of the most popular use cases on the Anvil Composable Subsystem is JupyterHub, an interactive, web-based data analysis toolkit.
- We provide a mechanism to integrate federated identity management via CILogon and user identity mapping via LDAP into Kubernetes-based Zero-to-Jupyterhub deployments.
- I-GUIDE is supporting integrative discovery through the analysis of large geospatial datasets via Jupyter and HPC.
- Our solution has been prototyped on Anvil to support the geospatial data analysis needs of I-GUIDE providing integration between the I-GUIDE hub in the Anvil Composable Subsystem and Anvil's GPFS storage.

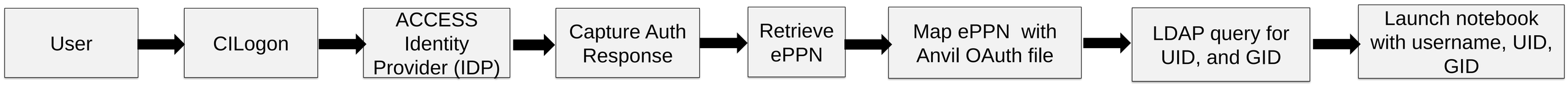
CILogon

I-GUIDE
Institute for Geospatial Understanding through an Integrative Discovery Environment

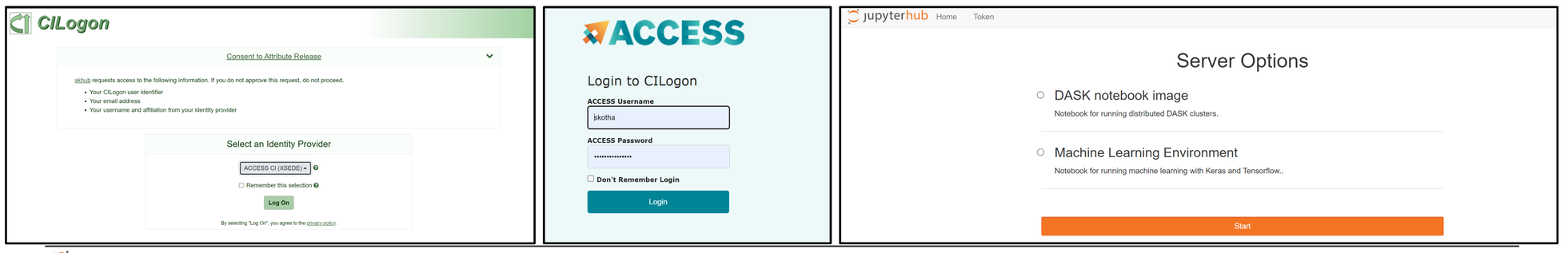
kubernetes

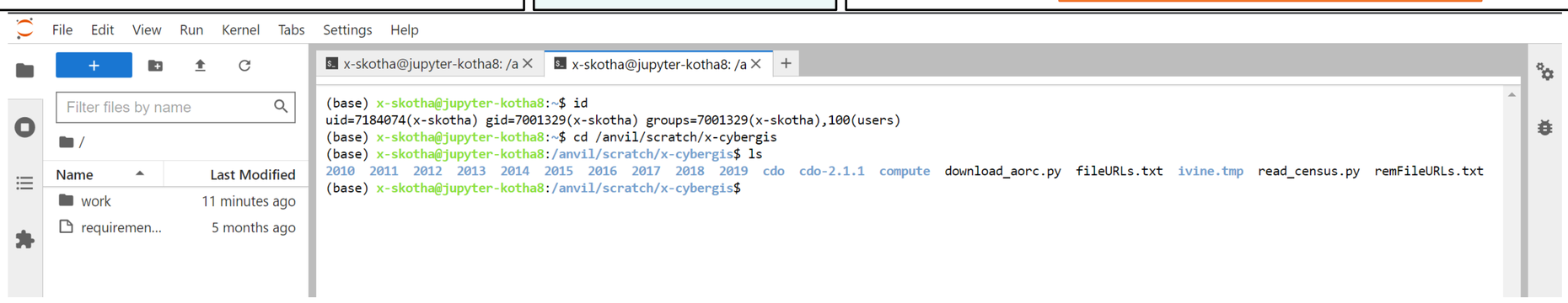
jupyterhub

Authentication Flowchart



Results





Future Work

Our solution allows users to run containers launched via JupyterHub instances as their Anvil uids. We plan to expand user identity mapping to be available to any container running on the Anvil's composable cloud. To accomplish this, we are investigating using Kyverno, a Kubernetes native policy management framework.

Acknowledgement

This work is supported by NSF OAC Project #2005632: Category I: Anvil - A National Composable Advanced Computational Resource for the Future of Science and Engineering

NSF

PURDUE UNIVERSITY

CERIAS