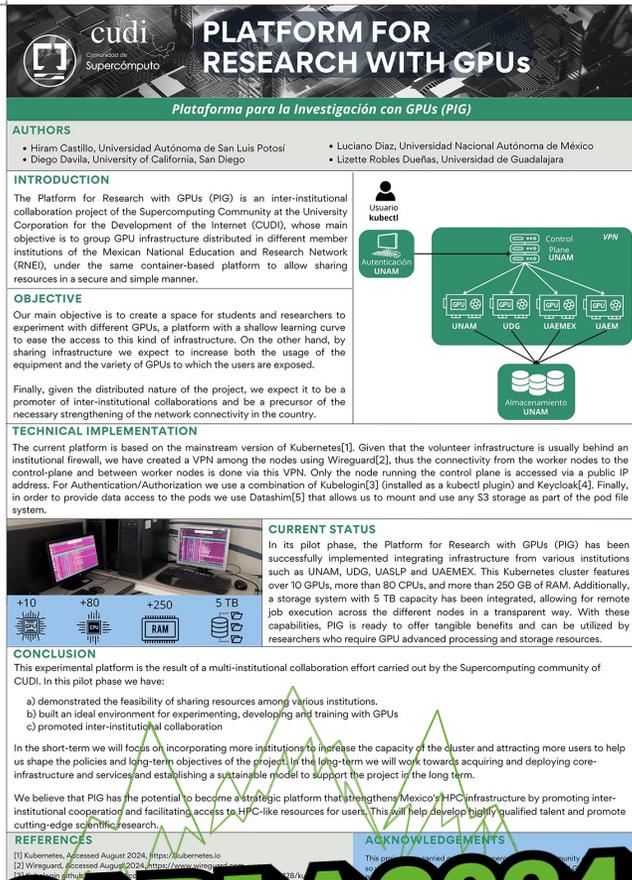


Internacional



cudi
Comunidad de Supercómputo

PLATFORM FOR RESEARCH WITH GPUS

Plataforma para la Investigación con GPUs (PIG)

AUTHORS

- Hiram Castillo, Universidad Autónoma de San Luis Potosí
- Diego Davila, University of California, San Diego
- Luciano Diaz, Universidad Nacional Autónoma de México
- Lizette Robles Dueñas, Universidad de Guadalajara

INTRODUCTION

The Platform for Research with GPUs (PIG) is an inter-institutional collaboration project of the Supercomputing Community at the University Corporation for the Development of the Internet (CUDI), whose main objective is to group GPU infrastructure distributed in different member institutions of the Mexican National Education and Research Network (RNEI), under the same container-based platform to allow sharing resources in a secure and simple manner.

OBJECTIVE

Our main objective is to create a space for students and researchers to experiment with different GPUs, a platform with a shallow learning curve to ease the access to this kind of infrastructure. On the other hand, by sharing infrastructure we expect to increase both the usage of the equipment and the variety of GPUs to which the users are exposed.

Finally, given the distributed nature of the project, we expect it to be a promoter of inter-institutional collaborations and be a precursor of the necessary strengthening of the network connectivity in the country.

TECHNICAL IMPLEMENTATION

The current platform is based on the mainstream version of Kubernetes[1]. Given that the volunteer infrastructure is usually behind an institutional firewall, we have created a VPN among the nodes using Wireguard[2], thus the connectivity from the worker nodes to the control-plane and between worker nodes is done via this VPN. Only the node running the control plane is accessed via a public IP address. For Authentication/Authorization we use a combination of Kubelogin[3] (installed as a kubectl plugin) and Keycloak[4]. Finally, in order to provide data access to the pods we use Datashim[5] that allows us to mount and use any S3 storage as part of the pod file system.

CURRENT STATUS

In its pilot phase, the Platform for Research with GPUs (PIG) has been successfully implemented integrating infrastructure from various institutions such as UNAM, UDG, UASLP and UAEMEX. This Kubernetes cluster features over 10 GPUs, more than 80 CPUs, and more than 250 GB of RAM. Additionally, a storage system with 5 TB capacity has been integrated, allowing for remote job execution across the different nodes in a transparent way. With these capabilities, PIG is ready to offer tangible benefits and can be utilized by researchers who require GPU advanced processing and storage resources.

CONCLUSION

This experimental platform is the result of a multi-institutional collaboration effort carried out by the Supercomputing community of CUDI. In this pilot phase we have:

- demonstrated the feasibility of sharing resources among various institutions.
- built an ideal environment for experimenting, developing and training with GPUs
- promoted inter-institutional collaboration

In the short-term we will focus on incorporating more institutions to increase the capacity of the cluster and attracting more users to help us shape the policies and long-term objectives of the project. In the long-term we will work towards acquiring and deploying core-infrastructure and services and establishing a sustainable model to support the project for the long term.

We believe that PIG has the potential to become a strategic platform that strengthens Mexico's HPC infrastructure by promoting inter-institutional cooperation and facilitating access to HPC-like resources for users. This will help develop highly qualified talent and promote cutting-edge scientific research.

REFERENCES

[1] Kubernetes, accessed August 2024, <https://kubernetes.io/>
 [2] Wireguard, accessed August 2024, <https://www.wireguard.com/>
 [3] Kubelogin, accessed August 2024, <https://github.com/kubelogin/kubelogin>
 [4] Keycloak, accessed August 2024, <https://www.keycloak.org/>
 [5] Datashim, accessed August 2024, <https://github.com/GoogleCloudPlatform/datashim>

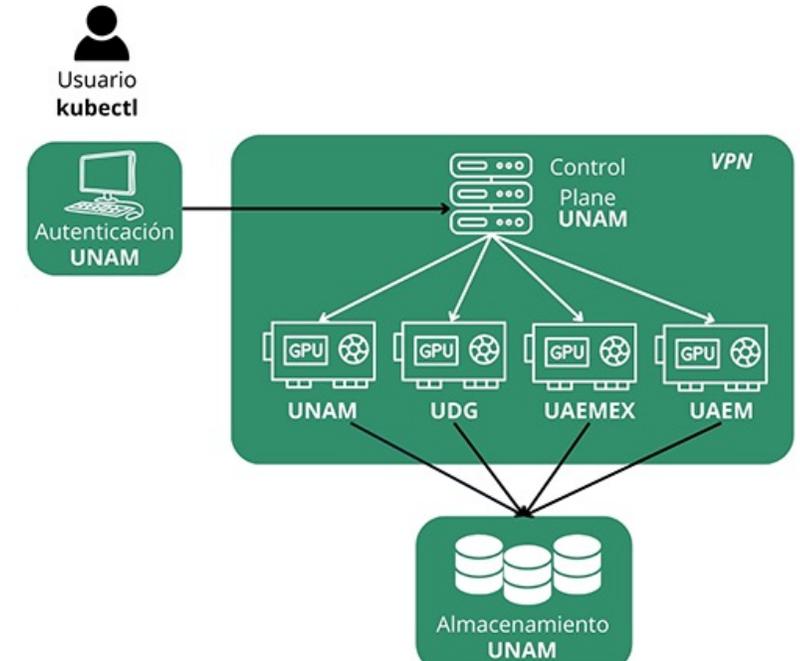
ACKNOWLEDGEMENTS

The platform provided by the Supercomputing Community at the University Corporation for the Development of the Internet (CUDI).

Coordinador: Lizette Robles - UDG

PIG, la Plataforma para la Investigación con GPUs, presente en CARLA2024

Los días 2 y 3 de octubre del presente año, se reunió en Santiago de Chile, la comunidad de Computación de Alto Rendimiento (HPC) en América Latina, con el objetivo de fomentar el crecimiento y fortalecer a la comunidad a través del intercambio y la difusión de nuevas ideas, técnicas e investigaciones en HPC y sus áreas de aplicaciones.



CARLA 2024
 LATIN AMERICA HIGH PERFORMANCE COMPUTING CONFERENCE
 SANTIAGO DE CHILE
 September 30 to October 4

Propuesta trabajo

- Continuar trabajando con los grupos de trabajo para impulsar proyectos relacionados con:
 - Cluster GPUs (PIG)
 - Cluster de HPC
 - Programa de capacitación
 - Actividades de divulgación
 - Actualización del inventario de la infraestructura instalada
 - Reconocimientos a los trabajos