SC21 Network Research Exhibition: Demonstration Abstract

StarLight DTN-as-a-Service and Kubernetes Integration for High-Performance Data Movement Support with Research Platforms

Se Young Yu, Jim Chen, Fei Yeh, Joe Mambretti
International Center for Advanced Internet Research - Northwestern University, young.yu, jim-chen, fyeh, j-mambretti@northwestern.edu

Abstract

DTN-as-a-Service focuses on moving large data in a cloud environment such as Kubernetes to improve the performance of the data movement over the high-performance networks. We implement cloud-native services for data movement within and among Kubernetes clouds through the DTN-as-a-Service framework, which sets up, optimizes, and monitors the underlying system and network. DTN-as-a-Service provides APIs to identify, examine and tune the underlying node for high-performance data movement in Kubernetes and enables data movement over a long-distance network. To map the big-data transfer workflow to a science workflow, a controller is implemented in Jupyter notebooks, a popular tool for data science.

Goals

1. StarLight DTN-as-a-Service focuses on moving large data in cloud environments, such as Kubernetes, to improve data transmission performance over high-performance networks.

2. We implement cloud-native services for data movement within and among Kubernetes clouds through DTN-as-a-service to set up, optimize, and monitor underlying systems and networks using Jupyter notebooks.

3. We demonstrate control of data movement between nodes in Kubernetes using cloud-native services implemented in DTN-as-a-Service to improve and analyze the performance over high-performance networks.

4. StarLight DTN-as-a-Service is designed to work with cloud environments specifically to optimize underlying resources in Kubernetes (compute, storage, memory, and network) in scalable service deployment and orchestration.

5. Related research includes cloud-native services, such as direct access to the low-level hardware from Kubernetes pods, network stack management, storage management, transfer protocol optimization, and monitoring of system and networks.

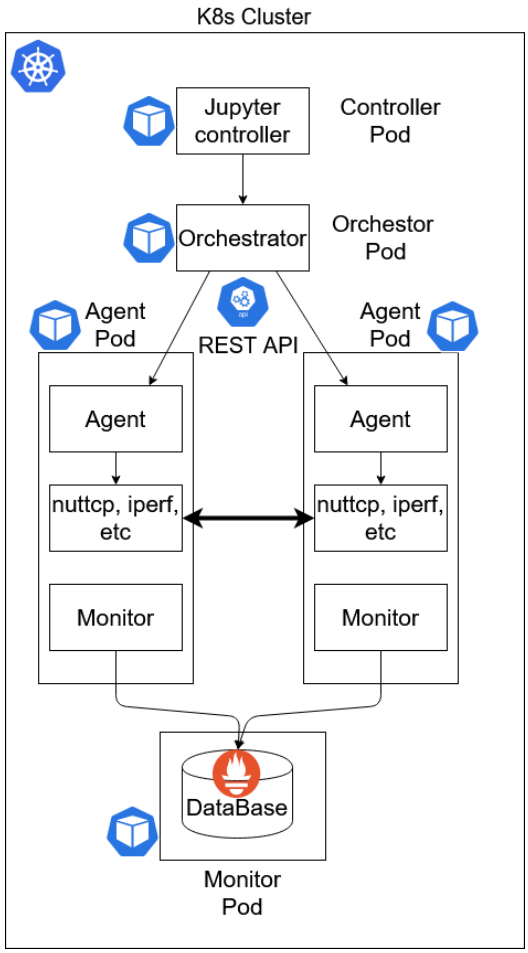

Figure 1. DTN-as-a-Service in Kubernetes

6. Enhancement includes additional capabilities of Kubernetes to DTN-as-a-Service software stack to control resources in the DTN for data movement using cloud-native services and Jupyter notebook and analysis of data movement through real-time monitoring of the resources.

Resources

Required resources from SCinet are use of some portion of the 1.2 Tbps path SCinet has been asked to provision from the StarLight facility in Chicago to the StarLight booth on the SC21 showfloor.

Involved Parties

- Se-Young Yu, iCAIR, young.yu@northwestern.edu
- Jim Chen, iCAIR, jim-chen@northwestern.edu
- Fei Yeh, iCAIR,fyeh@northwestern.edu
- Joe Mambretti, iCAIR, j-mambretti@northwestern.edu