SC22 Network Research Exhibition: Demonstration Abstract

NRE-003: StarLight DTN-as-a-Service and Kubernetes Integration for High-Performance Data Transport 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

For SC22, iCAIR will build on previous DTN-as-a-Service initiatives to extend those services to 400 Gbps capabilities. DTN-as-a-Service focuses on transporting large data across WANs and within a cloud environment such as Kubernetes to improve the data movement performance over high-performance networks. We implement cloud-native services for data transport 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 streamlines big data movement workflow using Jupyter controller, a popular data science tool, to identify, examine and tune the underlying DTNs for high-performance data movement in Kubernetes. In addition, it enables data movement over a long-distance network using different networking fabrics.

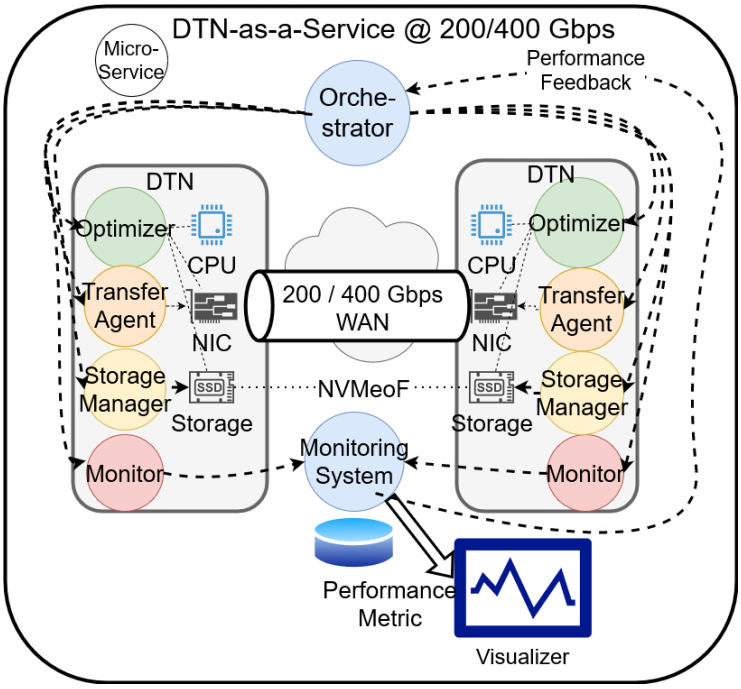
We achieved 194 Gbps End-to-End TCP throughput over an 88ms WAN path between StarLight and ESnet Testbed in Berkeley. We want to extend the testbed to include more sites and capabilities, such as dynamic path provisioning over loopbacks.

Goals

1. The StarLight Exchange DTN-as-a-Service focuses on moving large data in cloud environments, such as Kubernetes, on improving data transmission performance over high-performance networks. For SC22, these capabilities will be based on 400 Gbps switches and 400 Gbps servers using microservice architecture.

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

3. We demonstrate data movement control 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 systems and networks.

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.

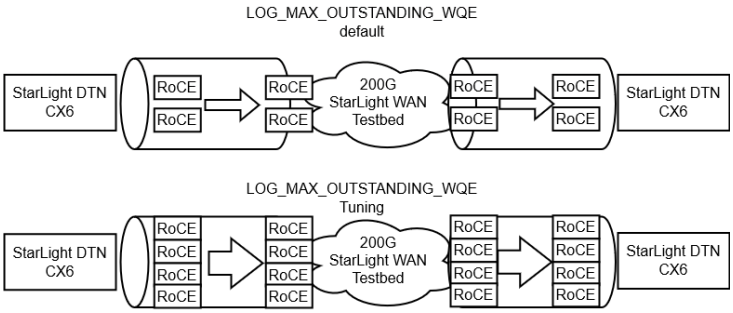


Figure 2. NIC tunning for 200G WAN services

Resources

Required resources from SCinet WAN are 1 Tbps E2E WAN services from the StarLight International/National Communications Exchange Facility in Chicago to the SC22 venue, between StarLight and the JBDT Facility in McLean, between the JBDT Facility and the SC22 venue and among all sites. In addition, another site utilized will be a 400 Gbps ESnet testbed at Berkeley connected to the StarLight Facility.

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