SC22 Network Research Exhibition: Demonstration Preliminary Abstract

[Optimizing Big Data Transfers using AI Strategies]

Gauravdeep Shami, Marc Lyonnais, Ciena, gshami,mlyonnai@ciena.com

Se-young Yu, Jim Chen, Joe Mambretti, iCair, young.yu, jim-chen, j-mambretti@northwestern.edu

Danial Ebling, Joe Breen, Jim Stewart, UETN, danial, jstewart@uen.org, Joe.Breen@utah.edu

Abstract

Massive scientific data flows are extremely time sensitive yet fragile as they are dependent on system capabilities and transient characteristics of the infrastructure. To achieve guaranteed high disk-to-disk throughput between end systems, the DTN hardware, OS parameters, software/orchestration stack, data transfer protocol and file management algorithm need to be customized as per use and hence no one size fits all. In this winning solution, called ODaaS: Optimized DTN as a Service, from the Data Mover Challenge presented in 2021, we propose a method which includes obtaining parameters pertaining to a data source, sink and a plurality of network elements and links configured along one or more data paths between the data source and data sink. The method then performs the step of automatically creating a high-bandwidth data transfer strategy for transferring a massive amount of data from the data source to the data sink based these extracted parameters.

Goals

The ODaaS solution aims to purport the following characteristics:

1. Ability to tune DTNs (OS, Network, storage) to an optimal configuration based on capabilities of the installed hardware.
2. Implement a data transfer protocol that reduces I/O overhead and directs data transfers to end process.
3. A data management entity that optimizes file sizes to achieve maximum throughput.
4. Security and identity management systems compatible widely in the industry.
5. Real-time predictive analytics and system monitoring capabilities that can advise and help fine-tune the transfer strategy.

Resources

The solution will be demonstrated using 100Gbps capable Data Transfer Nodes positioned on the Ciena Environment for Network Innovation in Ottawa, Hanover, iCair DTNS in Chicago and UETN DTNs in Salt Lake City. There is no request from Scinet at the moment.

Involved Parties

Demo will involve a number of parties such as Ciena, iCair, UETN as mentioned above.

- Gauravdeep Shami, Marc Lyonnais, Rodney Wilson, Scott Kohlert, Ciena Corporation, gshami,mlyonnai,rwilson,skohlert@ciena.com
- Se-young Yu, Jim Chen, Joe Mambretti, iCair, young.yu, jim-chen, j-mambretti@northwestern.edu
- Danial Ebling, Joe Breen, Jim Stewart, UETN, danial, jstewart@uen.org, Joe.Breen@utah.edu