

Applying P4 To Supporting Data Intensive Science Workflows On Large Scale Networks

Joe Mambretti, Jim Chen, Fei Yeh

International Center for Advanced Internet Research - Northwestern University

j-mambretti, jim-chen, fyeh@northwestern.edu

Abstract

[These demonstrations will show how P4 can be used to support large scale data intensive science workflows on high capacity high performance WANs and LANs. It has long been recognized that managing large scale data intensive science workflows on high performance LANs and WANs requires special programmable networking techniques. Recently, many of these techniques have been based on Software Defined Networking architecture and technology using the OpenFlow protocols. However, OpenFlow has multiple limitations, and, consequently, network researchers have begun to develop actual programming languages to provide for more agile, dynamic network services and capabilities. P4 is an emerging language for programmable networks that has a potential to address many issues related to supporting large scale data intensive science workflows on high capacity high performance WANs and LANs.]

Goals

1. These demonstrations will show how P4 (Programming Protocol-Independent Packet Processors), an open architecture, open source language for programming networks, can be used to support large scale data intensive science workflows on high capacity high performance WANs and LANs by introducing into the network a high degree of differentiation not previously possible.
2. Managing large scale data intensive science workflows on high performance LANs and WANs requires special programmable networking techniques
3. P4 is an emerging language for programmable networks that has a potential to address many issues related to supporting large scale data intensive science workflows on high capacity high performance WANs and LANs.
4. The P4 networking programming language, a domain specific language for network protocols P4 has many more capabilities than

OpenFlow, which is based on a limited set of check/match/action processes.

5. P4 programming enables the definition of headers (actual packet header and metadata), parsers (state machine and checksum field lists), and lookup tables (match/action). It allows for explicit descriptions of how specifically to define the way a particular forwarding plane processes packets.
6. P4 provides a vocabulary and sets of primitives that enable the development of actual programs that can be compiled and implemented on switches and/or NICs. P4 ASICs have been developed and are now appearing on the market.

Resources

Required resources from SCinet are 12*100 Gbps circuits from the StarLight facility in Chicago to the StarLight booth on the SC18 showfloor

Involved Parties

- Joe Mambretti, iCAIR, j-mambretti@northwestern.edu
- Jim Chen, iCAIR, jim-chen@northwestern.edu
- Fei Yeh, iCAIR, fyeh@northwestern.edu
- NSF Global Environment for Network Innovations Program Office (GPO)
- StarLight International/National Communications Exchange Facility and Consortium
- Metropolitan Research and Education Network (MREN)
- SCinet