

Experience with the OpenDaylight Controller in a multi-vendor 1Tbps network

Julian Bunn, Caltech. December 1 2014.

Introduction

During SC14 we tested the Hydrogen release of OpenDaylight, a community open-source software framework for controlling Software Defined Networks, particularly switchgear running the OpenFlow protocol.

Topology and Tests

The detailed hardware setup at SC14 is explained elsewhere on these pages. For the OpenDaylight tests we used Brocade MLXe16 switches at the Caltech, iCAIR and Vanderbilt booths, together with an Extreme switch collocated at the Caltech booth. The setup and interconnects are shown below:

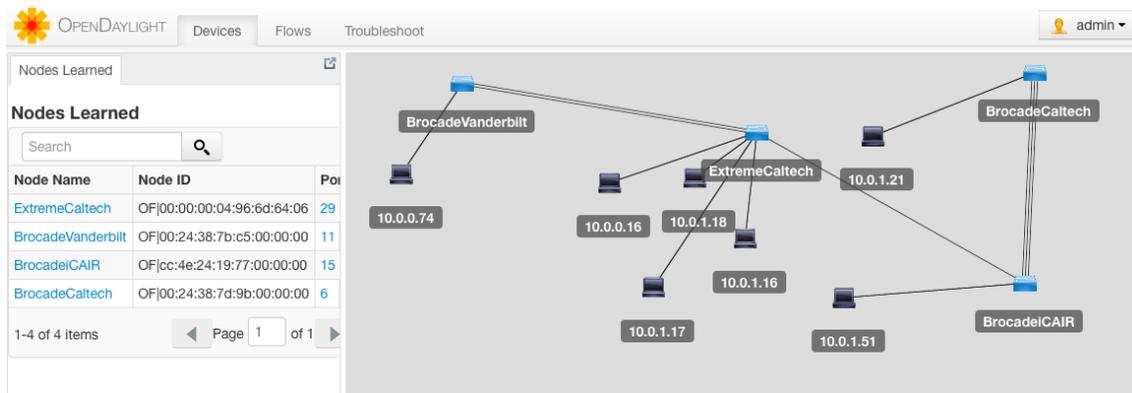


Figure 1 The OpenFlow switches and attached hosts used for the OpenDaylight tests at SC14.

The topology included a pair of 100Gbit links between the Vanderbilt Brocade and the Caltech Extreme, a single 100Gbit link between the Extreme and the iCAIR Brocade, and three 100 Gbit links between the iCAIR and Caltech Brocades. Each switch was configured to talk to the ODL controller, which was running in a dedicated Virtual Machine on a host at the Caltech booth.

The purpose of the OpenDaylight (ODL) tests was to evaluate the performance of plugins, developed at Caltech, that provide reactive intelligent routing of traffic between endpoints in the network. These plugins, Multipath and Multipath.Northbound¹ offer control over several routing strategies, as well as RESTful endpoints for control and monitoring of the network. In particular, the reactive mode of the Multipath plugin will respond to emergent traffic flows by installing flow rules in the switches along the dynamically chosen path for the flow. The path selection algorithms that were tested at SC14 included:

- Round Robin – each new traffic flow is assigned to the next possible path between source and destination
- Shortest Path – each flow is assigned to the shortest path (the least number of hops)

- Available Bandwidth – each flow is assigned to the path with the largest available bandwidth on the busiest hop
- Longest Path – each flow is assigned to the longest path (the greatest number of hops)
- Random Path – each flow is assigned to one of the available paths at random

(These are but a few of the available selectors provided – see the documentation for more details.)

To test the various selectors, the selector required was first configured using the Northbound API’s REST interface, and then traffic was started between hosts using either iperf or FDT, depending on the test. Each switch, on reception of a new frame from the source host, matches the frame against any existing flow entries in its switch tables. If the frame matches, the switch sends the packet on its way, as defined by the matching rule. If there is no match, the switch “punts” the packet to the ODL controller for inspection.

On reception of a punted packet/frame, the Multipath plugin decides whether to process the packet, depending on its protocol. For these tests ICMP and TCP packets in Ethernet V4 were processed – other punted packets were ignored. For both types of processed packets, the source and destination host address were extracted, together with their incoming outgoing switch ports. For TCP traffic, the source and destination TCP port numbers were optionally extracted for creating flow rules. (In most tests, the source TCP port was used to distinguish the flows.)

With the source and destination details for the punted packet being extracted, the Multipath plugin then used the current path selection algorithm to determine a path between the source and destination. Then, flow rules for each switch on the path were defined and accumulated in a list, before finally being sent to ODL’s “FlowManager” module for writing to the switches themselves. Extensive error checking is used to ensure that only complete, correct paths are left active in the switchgear.

The flow entries themselves used idle timeouts of either 10 seconds or 60 seconds (the latter being used early on during initial testing).

Results and Issues

The following screen shot shows a typical result during tests of the path selection algorithms:

```
"(OF|11@ExtremeCaltech->OF|2@BrocadeiCAIR)": " 0.0000 Gbits/sec",
"(OF|129@ExtremeCaltech->OF|290@BrocadeVanderbilt)": " 0.0000 Gbits/sec",
"(OF|139@ExtremeCaltech->OF|338@BrocadeVanderbilt)": " 0.0000 Gbits/sec",
"(OF|145@BrocadeiCAIR->OF|49@BrocadeCaltech)": " 0.0000 Gbits/sec",
"(OF|146@BrocadeiCAIR->OF|50@BrocadeCaltech)": " 0.0417 Gbits/sec",
"(OF|257@ExtremeCaltech->OF|290@BrocadeVanderbilt)": " 0.0476 Gbits/sec",
"(OF|267@ExtremeCaltech->OF|338@BrocadeVanderbilt)": " 0.0000 Gbits/sec",
"(OF|290@BrocadeVanderbilt->OF|129@ExtremeCaltech)": " 0.0000 Gbits/sec",
"(OF|290@BrocadeVanderbilt->OF|257@ExtremeCaltech)": " 12.6947 Gbits/sec",
"(OF|2@BrocadeCaltech->OF|98@BrocadeiCAIR)": " 12.3772 Gbits/sec",
"(OF|2@BrocadeiCAIR->OF|11@ExtremeCaltech)": " 0.0000 Gbits/sec",
"(OF|338@BrocadeVanderbilt->OF|139@ExtremeCaltech)": " 0.0000 Gbits/sec",
"(OF|338@BrocadeVanderbilt->OF|267@ExtremeCaltech)": " 23.7349 Gbits/sec",
"(OF|49@BrocadeCaltech->OF|145@BrocadeiCAIR)": " 12.3777 Gbits/sec",
"(OF|50@BrocadeCaltech->OF|146@BrocadeiCAIR)": " 12.3763 Gbits/sec",
"(OF|98@BrocadeiCAIR->OF|2@BrocadeCaltech)": " 0.0000 Gbits/sec"
```

Figure 2 Screen shot showing the Multipath.Northbound response to a request for DataRates during one of the SC14 tests

At this time, several FDT flows are in progress between Vanderbilt, iCAIR and Caltech following path selection with Multipath’s available bandwidth algorithm. In the following screen shot, traffic filling the 40Gbit/sec host adapters on a pair of nodes at iCAIR and Caltech is the result of an FDT flow set up using the Multipath system.

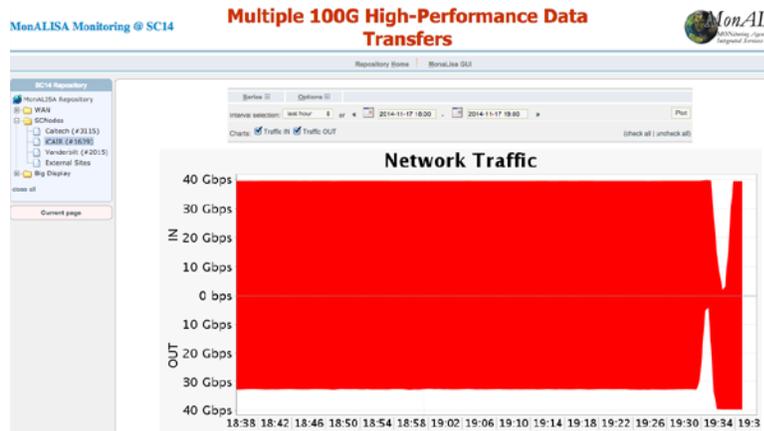


Figure 3 40Gbit/sec traffic flow between hosts at iCAIR and Caltech

There were several issues we faced during the tests:

1. The hardware configuration, and the switch ports assigned to OpenFlow control in the switches was dynamic during setup of the equipment, which made reliable testing challenging.
2. We observed some odd behavior of the firmware in the Extreme switch (e.g. punted packets matching existing flows, idle timeout values on flows not being respected) which we diagnosed with the help of Extreme technicians who were present at SC14 – we are most grateful for their help.
3. Patches were needed to the Brocade firmware to support some of the ODL switch commands, otherwise flow statistics were not available to ODL.
4. It was intended to include a remote Brocade (at the University of Victoria in BC) in the ODL topology, but problems with the level 1 link between there and New Orleans prevented this.

Helium

Towards the end of the SC14 event, we installed and successfully ran the Helium release of OpenDaylight. Although the Multipath modules were incompatible (and therefore not usable) with this release, we observed no problems. The following screenshot shows the topology – at this time the links between the switches were not configured for OpenFlow, and so are not visible.

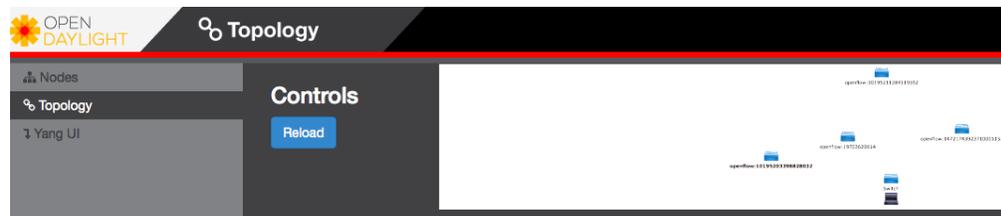


Figure 4 A screenshot showing the ODL topology at SC14, as detected by the OpenDaylight Helium installation.

ⁱ <http://pcbunn.cacr.caltech.edu/multipath/northbound/doc/index.html> and
<http://pcbunn.cacr.caltech.edu/multipath/doc/index.html>