

# OpenDaylight Performance Stress Test Report

## v1.3: Beryllium vs Boron

Intracom Telecom SDN/NFV Lab

[www.intracom-telecom.com](http://www.intracom-telecom.com) | [intracom-telecom-sdn.github.com](https://github.com/intracom-telecom-sdn) | [sdn-nfv@intracom-telecom.com](mailto:sdn-nfv@intracom-telecom.com)





# OpenDaylight Performance Stress Tests Report v1.3: Beryllium vs Boron

## Executive Summary

In this report we present comparative stress test results performed on OpenDaylight Beryllium SR0 controller against OpenDaylight Boron SR0. For these tests the NSTAT (Network Stress-Test Automation Toolkit) [1] testing platform and its external testing components (Multinet [2], OFTraf [3], MT-Cbench [4], nstat-nb-emulator [5]) have been used. The test cases presented in this report are identical to those presented in our previous performance report v1.2, so the interested reader is referred to [6] for comprehensive descriptions. As opposed to v1.2, in this report all test components have been executed within Docker containers [7] instead of KVM instances.

## Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>NSTAT Toolkit</b>                               | <b>3</b>  |
| <b>2</b> | <b>Experimental setup</b>                          | <b>4</b>  |
| <b>3</b> | <b>Switch scalability stress tests</b>             | <b>4</b>  |
| 3.1      | Idle Multinet switches                             | 4         |
| 3.1.1    | Test configuration                                 | 4         |
| 3.1.2    | Results  | 4         |
| 3.2      | Idle MT-Cbench switches                            | 5         |
| 3.2.1    | Test configuration                                 | 5         |
| 3.2.2    | Results  | 5         |
| 3.3      | Active Multinet switches                           | 5         |
| 3.3.1    | Test configuration                                 | 5         |
| 3.3.2    | Results  | 6         |
| 3.4      | Active MT-Cbench switches                          | 6         |
| 3.4.1    | Test configuration, "RPC" mode                     | 6         |
| 3.4.2    | Test configuration, "DataStore" mode               | 6         |
| 3.4.3    | Results  | 7         |
| <b>4</b> | <b>Stability tests</b>                             | <b>7</b>  |
| 4.1      | Idle Multinet switches                             | 7         |
| 4.1.1    | Test configuration                                 | 7         |
| 4.1.2    | Results  | 8         |
| 4.2      | Active MT-Cbench switches                          | 8         |
| 4.2.1    | Test configuration, "DataStore" mode               | 9         |
| 4.2.2    | Results  | 9         |
| 4.2.3    | Test configuration, "RPC" mode                     | 9         |
| 4.2.4    | Results  | 10        |
| <b>5</b> | <b>Flow scalability tests</b>                      | <b>10</b> |
| 5.1      | Test configuration                                 | 11        |
| 5.2      | Results  | 12        |
| 5.2.1    | Add controller time/rate                           | 12        |
| 5.2.2    | End-to-end flows installation controller time/rate | 12        |
| <b>6</b> | <b>Contributors</b>                                | <b>12</b> |
|          | <b>References</b>                                  | <b>12</b> |

## 1. NSTAT Toolkit

The NSTAT toolkit follows a modular and distributed architecture. With the term modular we mean that the core application works as an orchestrator that coordinates the testing

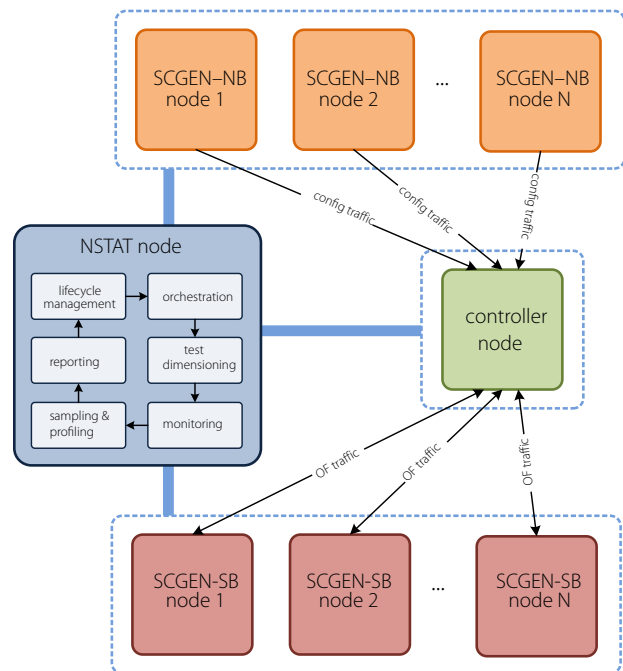
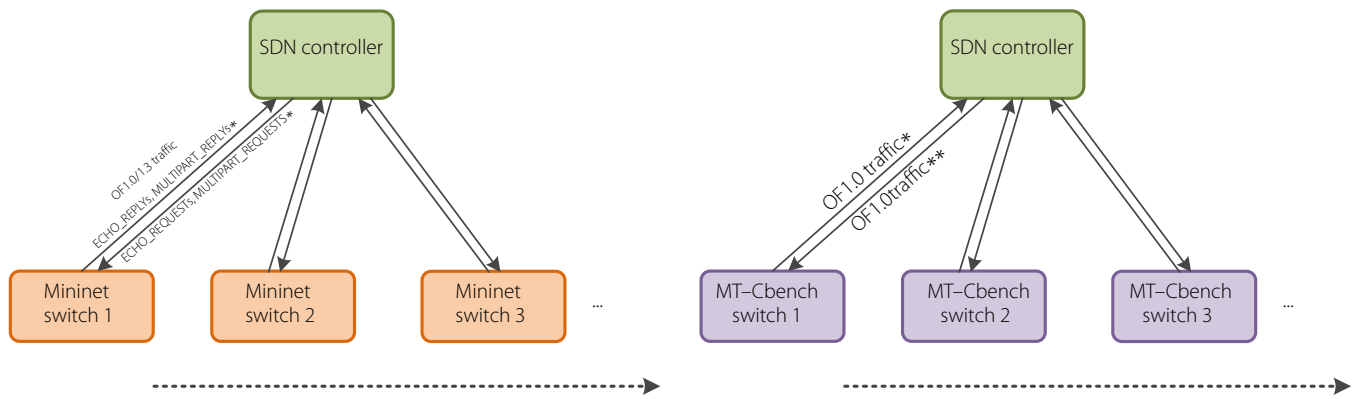


Fig. 1: NSTAT architecture

process. It controls the lifecycle of different testing components and also the coordination and lifecycle management of the testing subject, the OpenDaylight controller. With the term distributed, we mean that each component, controlled by NSTAT, can either run on the same node or on different nodes, which can be either physical or virtual machines. In the latest implementation of NSTAT we introduced the use of Docker containers [7].

With the use of containers<sup>1</sup> we can isolate the separate components and their use of resources. In older versions of NSTAT we were using CPU affinity [9] to achieve this resource isolation. In Fig.1 NSTAT toolkit architecture is depicted.

<sup>1</sup>The provisioning of docker containers along with their interconnection, is achieved with the use of 1) the docker-compose provisioning tool and 2) the pre-built docker images which are present at docker hub [8]. All containers were running on the same physical server.



\* the vast majority of packets

\*, \*\* a few messages during initialization, idle during main execution.

(a) Multinet switches sit idle during main operation. ECHO and MULTIPART messages dominate the traffic being exchanged between the controller and the switches.

(b) MT-Cbench switches sit idle during main operation, without sending or receiving any kind of messages.

Fig. 2: Representation of a switch scalability test with idle Multinet and MT-Cbench switches.

Table 1: Stress tests experimental setup.

|                            |  |
|----------------------------|--|
| Host operating system      | Centos 7, kernel 3.10.0                                  |
| Nodes type                 | Docker containers (Docker version 1.12.1, build 23cf638) |
| Container OS               | Ubuntu 14.04   |
| Physical Server platform   | Dell R720  |
| Processor model            | Intel Xeon CPU E5-2640 v2 @ 2.00GHz                      |
| Total system CPUs          | 32   |
| CPUs configuration         | 2 sockets × 8 cores/socket × 2 HW-threads/core @ 2.00GHz |
| Main memory                | 256 GB, 1.6 GHz RDIMM                                    |
| SDN Controllers under test | OpenDaylight Beryllium (SR1), OpenDaylight Boron (SR0)   |
| Controller JVM options     | -Xmx8G, -Xms8G, -XX:+UseG1GC, -XX:MaxPermSize=8G         |
| Multinet OVS version       | 2.3.0  |

## 2. Experimental setup

Details of the experimental setup are presented in Table 1.

## 3. Switch scalability stress tests

In switch scalability tests we test controller towards different scales of OpenFlow switches networks. In order to create these networks we use either MT-Cbench [4] or Multinet [2]. MT-Cbench generates OpenFlow traffic emulating a “fake” OpenFlow v1.0 switch topology. Multinet utilizes Mininet [10] and OpenVSwitch v2.3.0 [11] to emulate distributed OpenFlow v1.3 switch topologies.

In our stress tests we have experimented with topology switches operating in two modes, **idle** and **active** mode: switches in idle mode do not initiate any traffic to the controller, but rather respond to messages sent by it. Switches in active mode consistently initiate traffic to the controller, in the form of PACKET\_IN messages. In most stress tests, MT-Cbench and Multinet switches operate both in active and idle modes.

For more details regarding the tests setup, the reader should refer to the NSTAT: OpenDaylight Performance Stress Test Re-

port v1.2, [6]

### 3.1 Idle Multinet switches

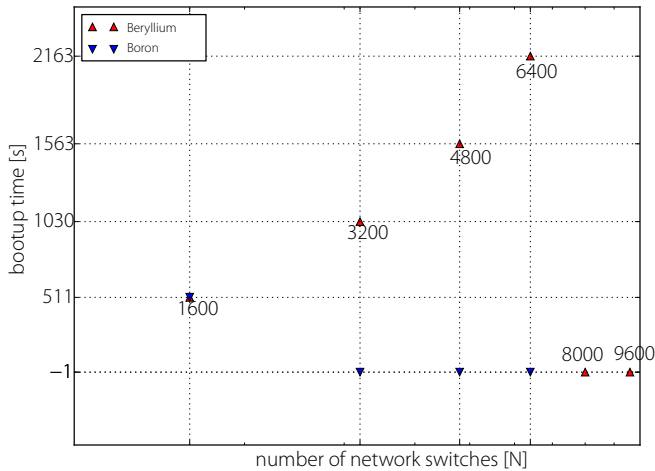
#### 3.1.1 Test configuration

- topology types: “Linear”, “Disconnected”
- topology size per worker node: 100, 200, 300, 400, 500, 600
- number of worker nodes: 16
- group size: 1, 5
- group delay: 5000ms
- persistence: enabled

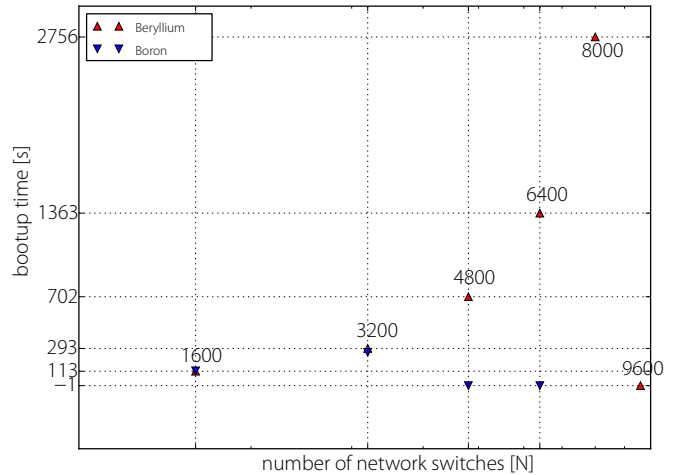
In this case, the total topology size is equal to 1600, 3200, 4800, 6400, 8000, 9600.

#### 3.1.2 Results

Results from this series of tests are presented in Fig.3(a), 3(b) for a “Linear” topology and Fig.4(a), 4(b) for a “Disconnected” topology respectively.

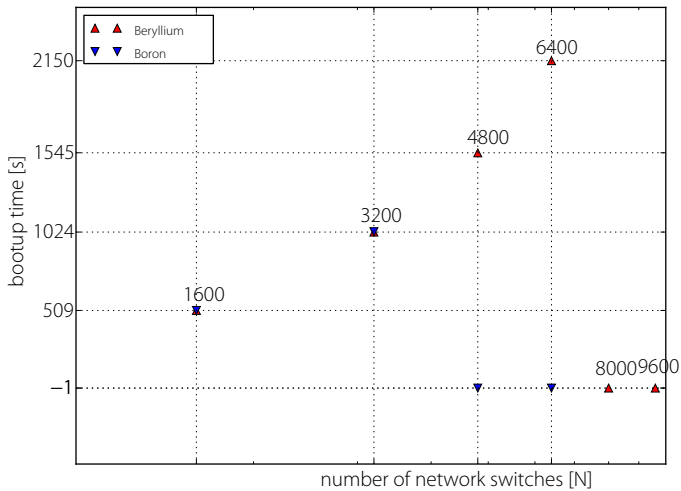


(a) Group size: 1, topology type: Linear

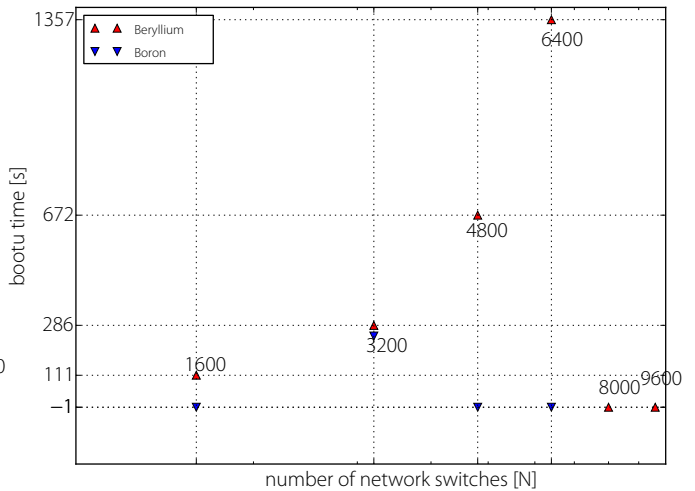


(b) Group size: 5, topology type: Linear

**Fig. 3:** Switch scalability stress test results for idle Multinet switches. Network topology size from 1600→9600 switches. Topology type: Linear. Boot-up time is forced to -1 when switch discovery fails.



(a) Group size: 1, topology type: Disconnected.



(b) Group size: 5, topology type: Disconnected.

**Fig. 4:** Switch scalability stress test results for idle Multinet switches. Network size scales from 1600→9600 switches. Topology type: Disconnected. Boot-up time is forced to -1 when switch discovery fails.

## 3.2 Idle MT-Cbench switches

### 3.2.1 Test configuration

- controller: "RPC" mode
- generator: MT-Cbench, latency mode
- number of MT-Cbench threads: 1, 2, 4, 8, 16, 20, 30, 40, 50, 60, 70, 80, 90, 100 threads.
- topology size per MT-Cbench thread: 50 switches
- group delay: 500, 1000, 2000, 4000, 8000, 16000.
- persistence: enabled

In this case switch topology size is equal to: 50, 100, 200, 300, 400, 800, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000.

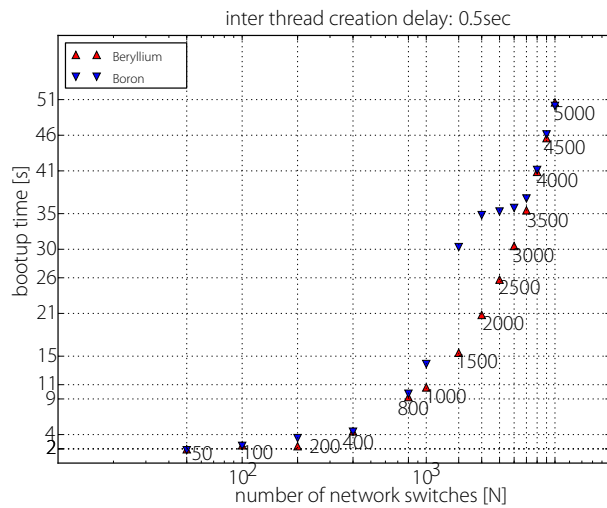
### 3.2.2 Results

Results of this test are presented in Fig.5(a), 5(b), 6(a), 6(b).

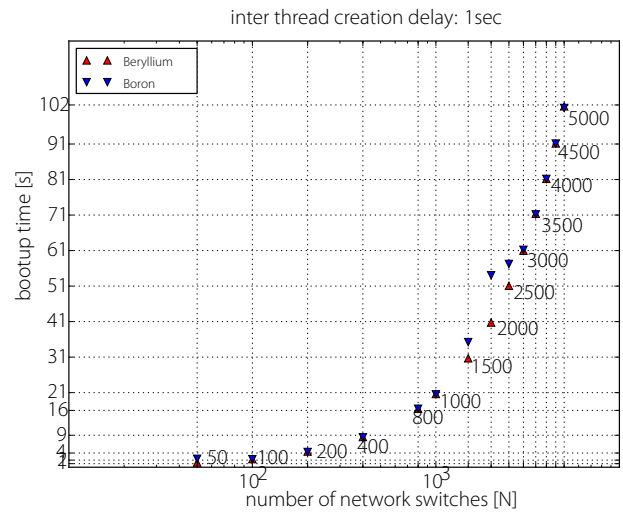
## 3.3 Active Multinet switches

### 3.3.1 Test configuration

- controller: with l2switch plugin installed, configured to respond with mac-to-mac FLOW\_MODs to PACKET\_IN messages with ARP payload [12]
- topology size per worker node: 12, 25, 50, 100, 200, 300
- number of worker nodes: 16
- group size: 1
- group delay: 3000ms
- topology type: "Linear"

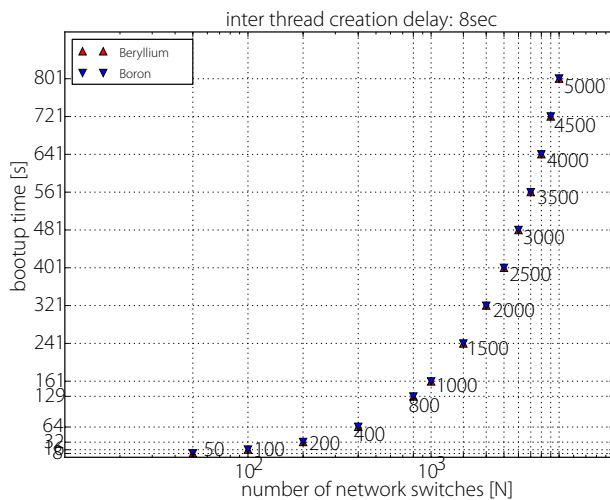


(a) group delay: 0.5s.

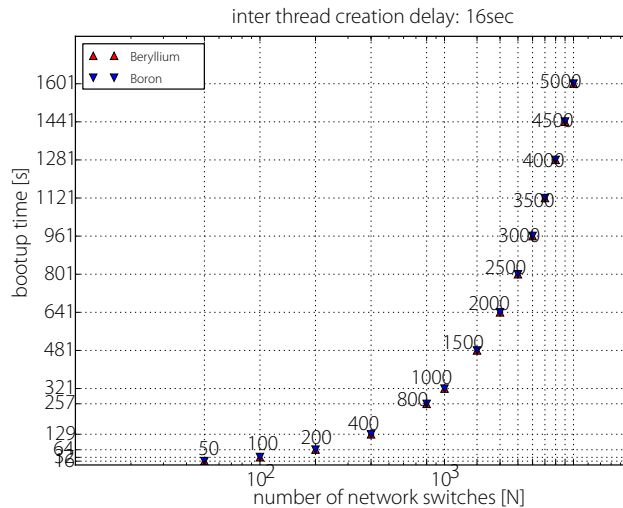


(b) group delay: 1s.

**Fig. 5:** Switch scalability stress test results with idle MT-Cbench switches. Topology size scales from 50 → 5000 switches.



(a) group delay: 8s.



(b) group delay: 16s.

**Fig. 6:** Switch scalability stress test results with idle MT-Cbench switches. Topology size scales from 50 → 5000 switches.

- hosts per switch: 2
- traffic generation interval: 60000ms
- PACKET\_IN transmission delay: 500ms
- persistence: disabled

- generator: MT-Cbench, latency mode
- number of MT-Cbench threads: 1, 2, 4, 8, 16, 20, 40, 60, 80, 100 threads,
- topology size per MT-Cbench thread: 50 switches
- group delay: 15s
- traffic generation interval: 20s
- persistence: enabled

Switch topology size scales as follows: 192, 400, 800, 1600, 3200, 4800.

### 3.3.2 Results

Results of this test are presented in Fig.8.

## 3.4 Active MT-Cbench switches

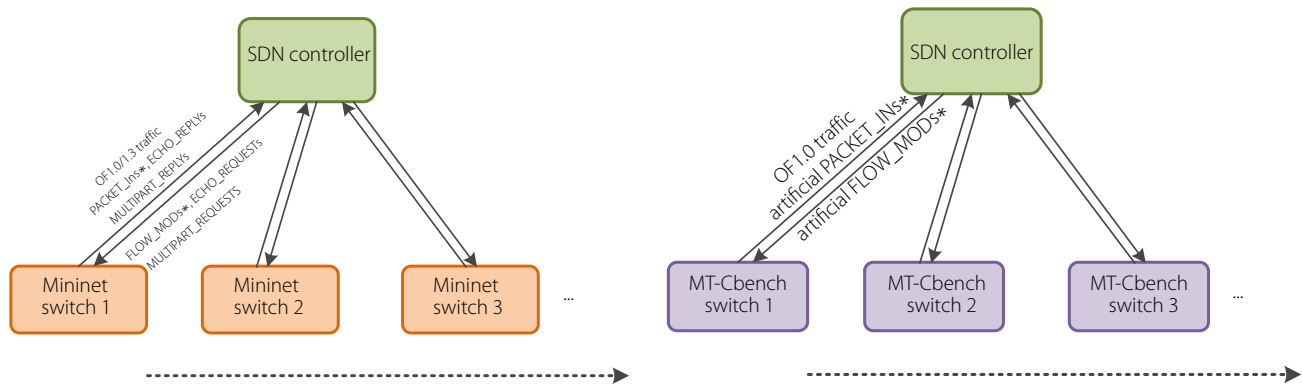
### 3.4.1 Test configuration, "RPC" mode

- controller: "RPC" mode

### 3.4.2 Test configuration, "DataStore" mode

- controller: "DataStore" mode
- generator: MT-Cbench, latency mode,
- number of MT-Cbench threads: 1, 2, 4, 8, 16, 20, 40, 60, 80, 100 threads,

In this case, the total topology size is equal to 50, 100, 200, 400, 800, 1000, 2000, 3000, 4000, 5000.



\* the vast majority of packets

(a) The switches send OF1.3 PACKET\_IN messages to the controller, which replies with OF1.3 FLOW\_MODs.

\* the vast majority of packets

(b) The switches send artificial OF1.0 PACKET\_IN messages to the controller, which replies with also artificial OF1.0 FLOW\_MOD messages.

Fig. 7: Representation of switch scalability stress test with active (a) Multinet and (b) MT-Cbench switches.

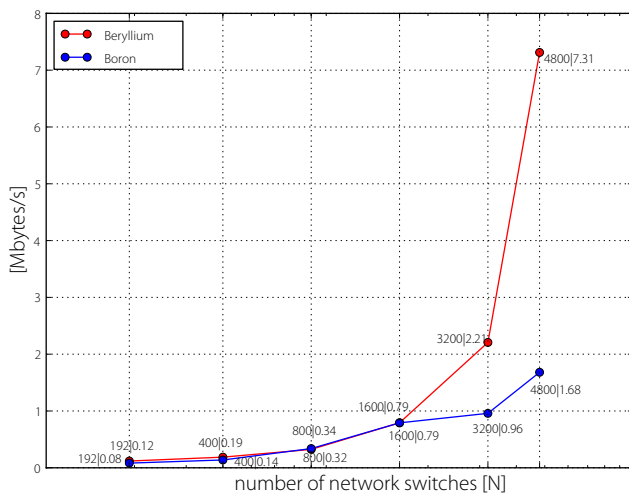


Fig. 8: Switch scalability stress test results with active Multinet switches. Comparison analysis of controller throughput variation Mbytes/s vs number of network switches N. Topology size scales from 192 → 4800 switches.

- topology size per MT-Cbench thread: 50 switches
- group delay: 15s
- traffic generation interval: 20s
- persistence: enabled

In this case, the total topology size is equal to 50, 100, 200, 300, 400, 800, 1000, 2000, 3000, 4000, 5000.

### 3.4.3 Results

Results for test configurations defined in sections 3.4.1, 3.4.2 are presented in Figs.9(a), 9(b) respectively.

## 4. Stability tests

Stability tests explore how controller throughput behaves in a large time window with a fixed topology connected to it, Fig.10(a), 10(b). The goal is to detect performance fluctuations over time.

The controller accepts a standard rate of incoming traffic and its response throughput is being sampled periodically. NSTAT reports these samples in a time series.

### 4.1 Idle Multinet switches

The purpose of this test is to investigate the stability of the controller to serve standard traffic requirements of a large scale Multinet topology of idle switches.

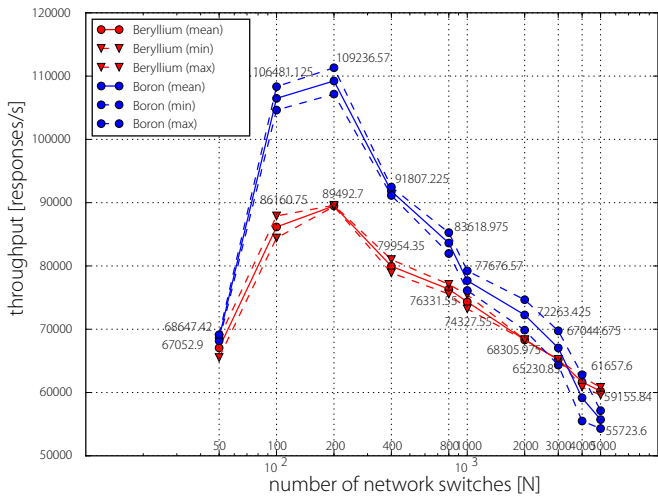
During main test execution Multinet switches respond to ECHO and MULTIPART messages sent by the controller at regular intervals. These types of messages dominate the total traffic volume during execution.

NSTAT uses the oftraf [3] to measure the **outgoing traffic of the controller**. The metrics presented for this case are the OpenFlow packets and bytes collected by NSTAT every second, Fig.15.

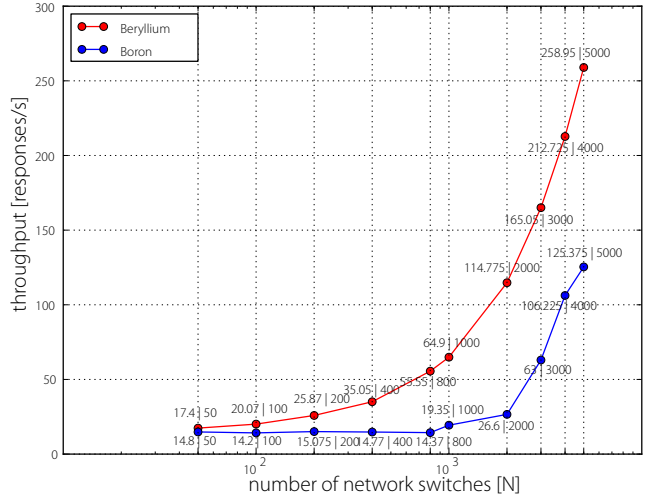
In order to push the controller performance to its limits, the controller is executed on the bare metal and Multinet on a set of interconnected virtual machines

#### 4.1.1 Test configuration

- topology size per worker node: 200
- number of worker nodes: 16
- group size: 1
- group delay: 2000ms
- topology type: "Linear"
- hosts per switch: 1

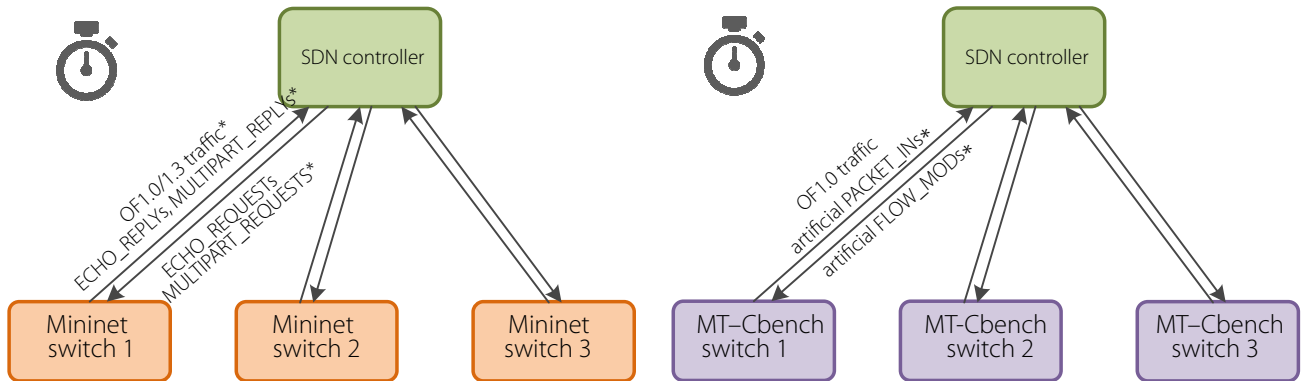


(a) OpenDaylight running in RPC mode.



(b) OpenDaylight running in DS mode.

**Fig. 9:** Switch scalability stress test results with active MT-Cbench switches. Comparison analysis of controller throughput variation [responses/s] vs number of network switches N with OpenDaylight running both on RPC and DataStore mode.



\* the vast majority of packets

(a) Representation of switch stability stress test with idle Multinet switches.

\* the vast majority of packets

(b) Representation of switch stability stress test with active MT-Cbench switches.

**Fig. 10:** Representation of switch stability stress test with idle Multinet (a) and active MT-Cbench switches (b). The controller accepts a standard rate of incoming traffic and its response throughput is being sampled periodically.

- period between samples: 10s
- number of samples: 4320
- persistence: enabled
- total running time: 12h

In this case switch topology size is equal to 3200.

#### 4.1.2 Results

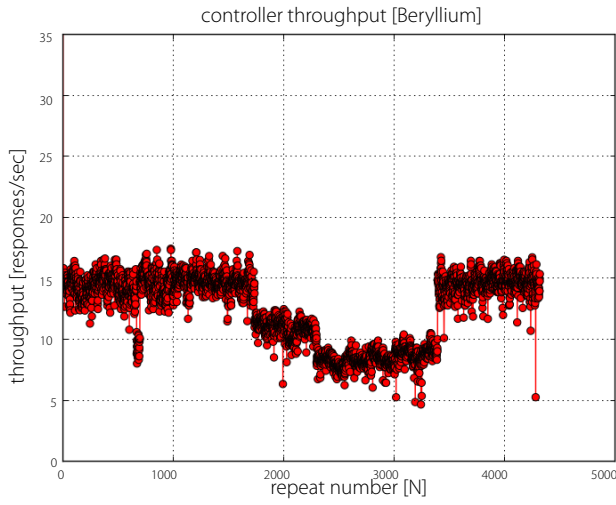
The results of this test are presented in Fig.15

#### 4.2 Active MT-Cbench switches

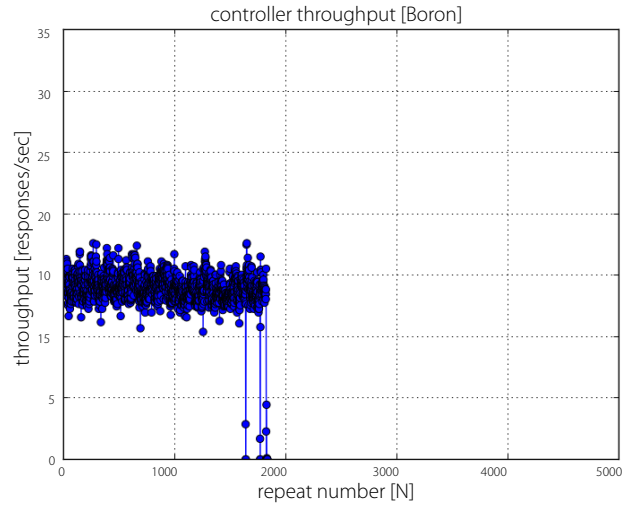
In this series of experiments NSTAT uses a fixed topology of active MT-Cbench switches to generate traffic for a large time window. The switches send artificial OF1.0 PACKET\_IN messages at a fixed rate to the controller, which replies with also artificial OF1.0 FLOW\_MOD messages; these message types dominate the traffic exchanged between the switches and the controller. We evaluate the controller both in "RPC" and "DataStore" modes.

In order to push the controller performance to its limits, all test nodes (controller, MT-Cbench) were executed on bare metal. To isolate the nodes from each other, the CPU shares feature



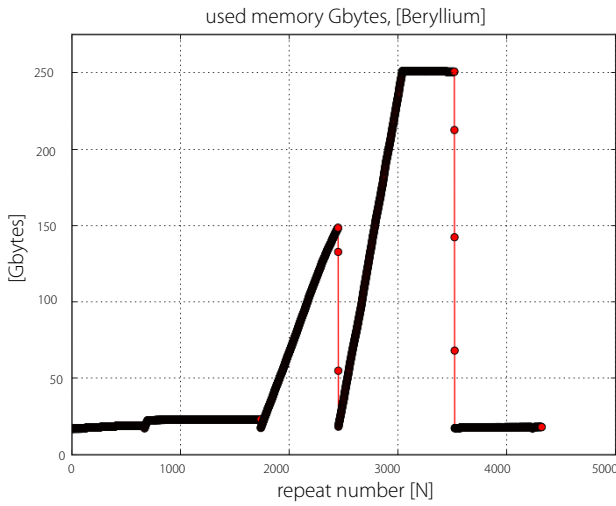


(a)

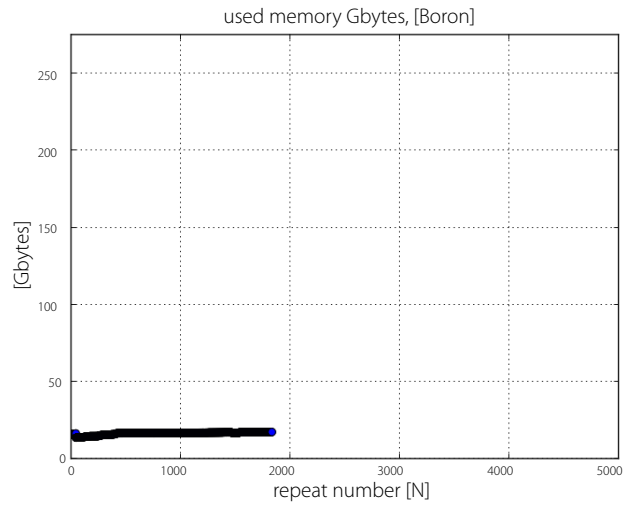


(b)

**Fig. 11:** Controller stability stress test with idle MT-Cbench switches. Throughput comparison analysis between OpenDaylight Beryllium and Boron versions. OpenDaylight running in "DS" mode.



(a)



(b)

**Fig. 12:** Controller stability stress test with idle MT-Cbench switches. Throughput comparison analysis between OpenDaylight Beryllium and Boron versions. OpenDaylight running in "DS" mode.

of NSTAT was used.

#### 4.2.1 Test configuration, "DataStore" mode

- controller: "DataStore" mode
- generator: MT-Cbench, latency mode
- number of MT-Cbench threads: 10
- topology size per MT-Cbench thread: 50 switches
- group delay: 8s
- number of samples: 4320
- period between samples: 10s
- persistence: enabled
- total running time: 12h

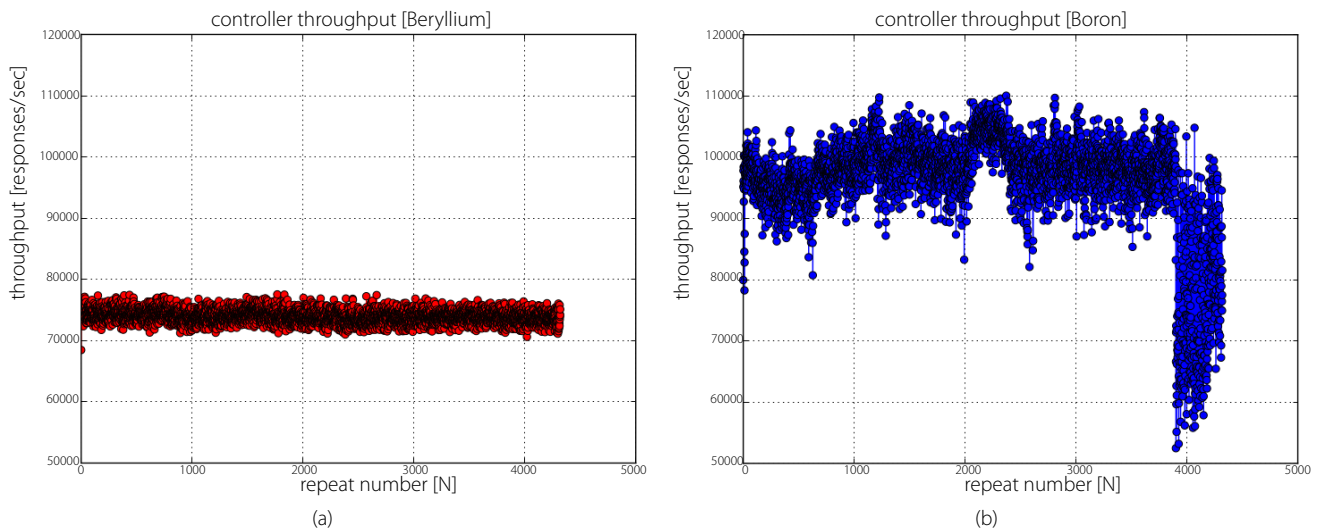
In this case, the total topology size is equal to 500 switches.

#### 4.2.2 Results

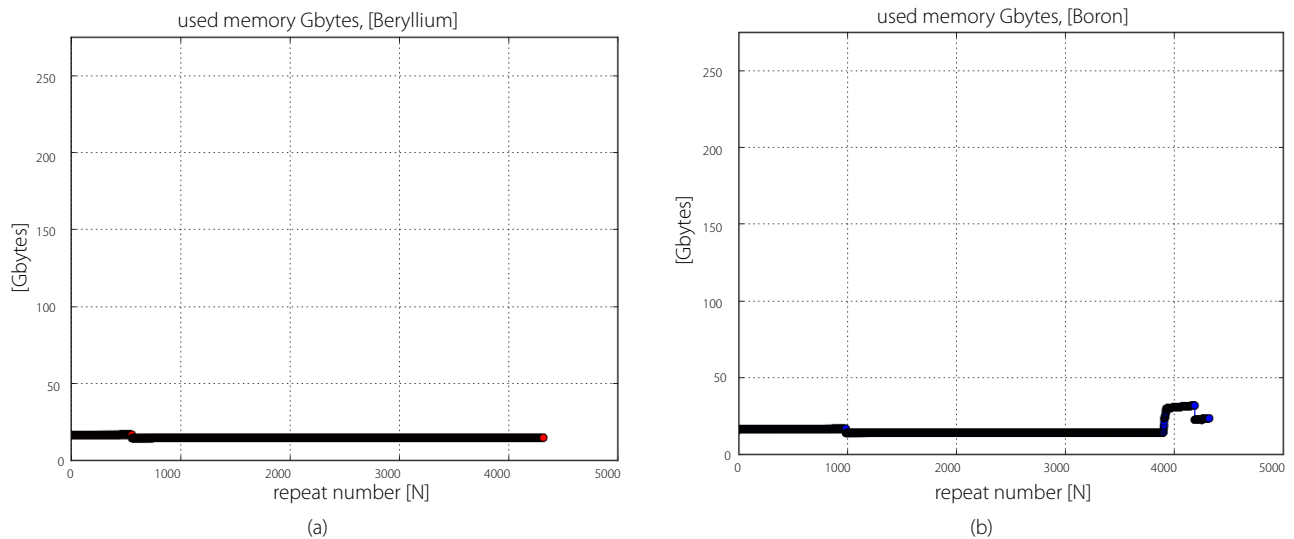
The results of this test are presented in Fig.11, 12.

#### 4.2.3 Test configuration, "RPC" mode

- controller: "RPC" mode
- generator: MT-Cbench, latency mode
- number of MT-Cbench threads: 10
- topology size per MT-Cbench thread: 50 switches
- group delay: 8s
- number of samples: 4320,



**Fig. 13:** Controller stability stress test with idle MT-Cbench switches. Throughput comparison analysis between OpenDaylight Beryllium and Boron versions. OpenDaylight running in "RPC" mode.



**Fig. 14:** Controller stability stress test with idle MT-Cbench switches. Throughput comparison analysis between OpenDaylight Beryllium and Boron versions. OpenDaylight running in "RPC" mode.

- period between samples: 10s
- persistence: enabled
- total running time: 12h

In this case, the total topology size is equal to 500 switches.

#### 4.2.4 Results

The results of this test are presented in Fig.13, 14.

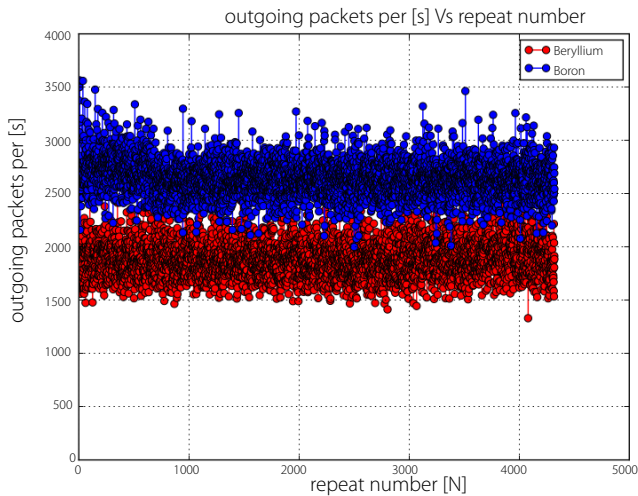
### 5. Flow scalability tests

With flow scalability stress tests we try to investigate both capacity and timing aspects of flows installation via the control-

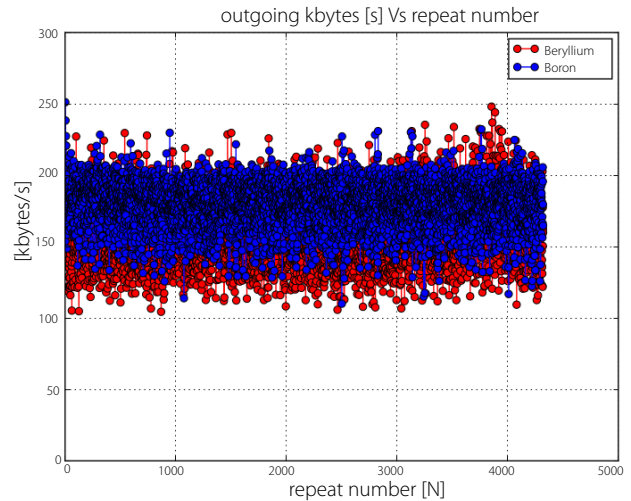
ler NB (RESTCONF) interface, Fig.16.

This test uses the NorthBound flow emulator [5] to create flows in a scalable and configurable manner (number of flows, delay between flows, number of flow writer threads). The flows are being written to the controller configuration data store via its NB interface, and then forwarded to an underlying Multinet topology as flow modifications, where they are distributed into switches in a balanced fashion.

The test verifies the success or failure of flow operations via the controller's operational data store. An experiment is considered successful when all flows have been installed on switches and have been reflected in the operational data store of the controller. If not all of them have become visible in the

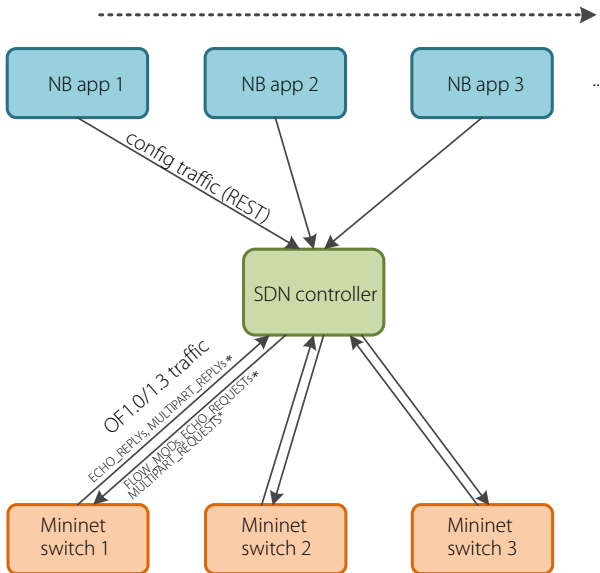


(a) Comparison analysis for outgoing packets/s between OpenDaylight Beryllium and Boron releases.



(b) Comparison analysis for outgoing kbytes/s between OpenDaylight Beryllium and Boron releases.

**Fig. 15:** Controller 12-hour stability stress test with idle Multinet switches.



\* the vast majority of packets

**Fig. 16:** Representation of flow scalability stress test. An increasing number of NB clients (NB app<sub>j</sub>, j=1,2,...) install flows on the switches of an underlying OpenFlow switch topology.

data store within a certain deadline after the last update, the experiment is considered failed.

Intuitively, this test emulates a scenario where multiple NB applications, each controlling a subset of the topology<sup>2</sup>, send simultaneously flows to their switches via the controller's NB interface at a configurable rate.

<sup>2</sup>subsets are non-overlapping and equally sized

The metrics measured in this test are:

- **Add controller time** ( $t_{add}$ ): the time needed for all requests to be sent and their response to be received (as in [9]).
- **Add controller rate** ( $r_{add}$ ):  $r_{add} = N / t_{add}$ , where N the aggregate number of flows to be installed by worker threads.
- **End-to-end installation time** ( $t_{e2e}$ ): the time from the first flow installation request until all flows have been installed and become visible in the operational data store.
- **End-to-end installation rate** ( $r_{e2e}$ ):  $r_{e2e} = N / t_{e2e}$

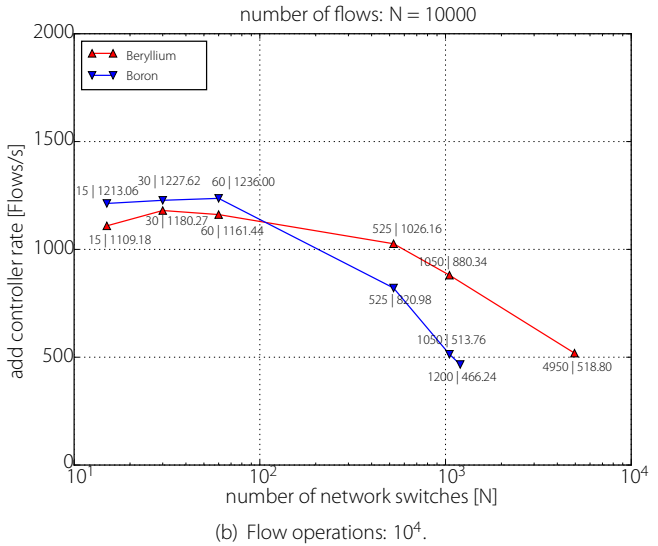
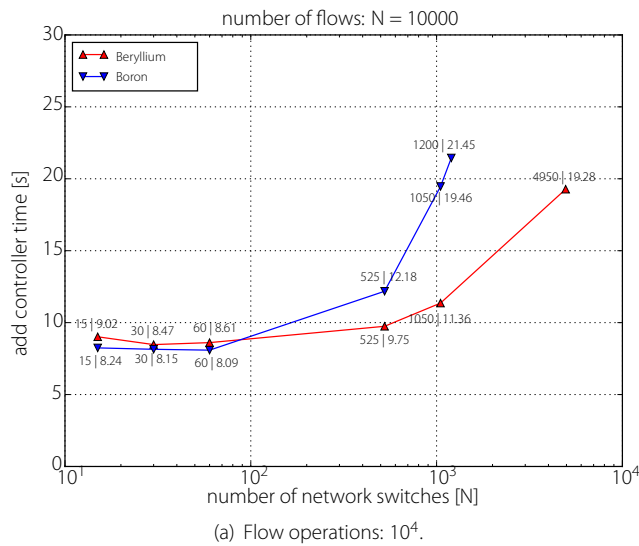
In this test, Multinet switches operate in idle mode, without initiating any traffic apart from the MULTIPART and ECHO messages with which they reply to controller's requests at regular intervals.

### 5.1 Test configuration

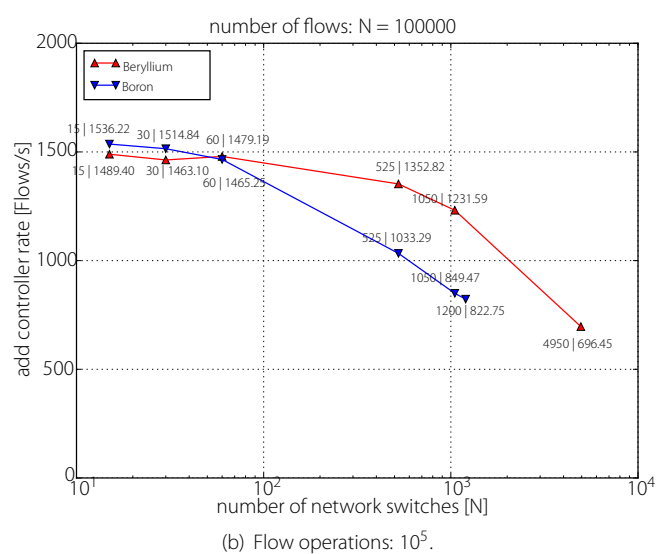
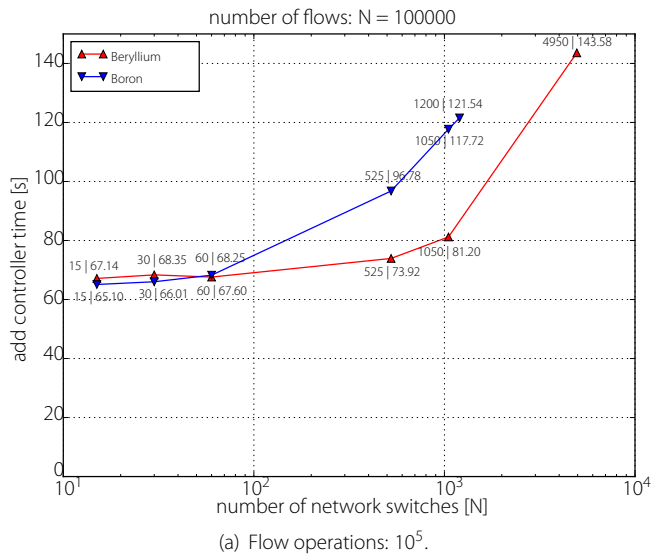
For both Lithium and Beryllium we used the following setup

- topology size per worker node: 1, 2, 4, 35, 70, 330.
- number of worker nodes: 15
- group size: 1
- group delay: 3000ms
- topology type: "Linear"
- hosts per switch: 1
- total flows to be added: 1K, 10K, 100K, 1M
- flow creation delay: 0ms
- flow worker threads: 5
- persistence: disabled

In this case switch topology size is equal to: 15, 30, 60, 525, 1050, 4950.



**Fig. 17:** Flow scalability stress test result. Comparison performance for add controller time/rate vs number of switches for various numbers of flow operations N. Add controller time is forced to -1 when test fails. Add controller time/rate vs number of switches for N=10<sup>4</sup> flow operations.



**Fig. 18:** Flow scalability stress test result. Comparison performance for add controller time/rate vs number of switches for various numbers of flow operations N. Add controller time is forced to -1 when test fails. Add controller time/rate vs number of switches for N=10<sup>5</sup> flow operations.

## 5.2 Results

### 5.2.1 Add controller time/rate

The results of this test are presented in Figs.17, 18, 19.

### 5.2.2 End-to-end flows installation controller time/rate

The results of this experiment are presented in Figs.20, 21.

## 6. Contributors

- Nikos Anastopoulos
- Panagiotis Georgiopoulos

- Konstantinos Papadopoulos
- Thomas Sounapoglou

## References

- [1] "NSTAT: Network Stress Test Automation Toolkit." <https://github.com/intracom-telecom-sdn/nstat>.
- [2] "Multinet: Large-scale SDN emulator based on Mininet." <https://github.com/intracom-telecom-sdn/multinet>.
- [3] "OFTraf: pcap-based, RESTful OpenFlow traffic monitor." <https://github.com/intracom-telecom-sdn/oftraf>.

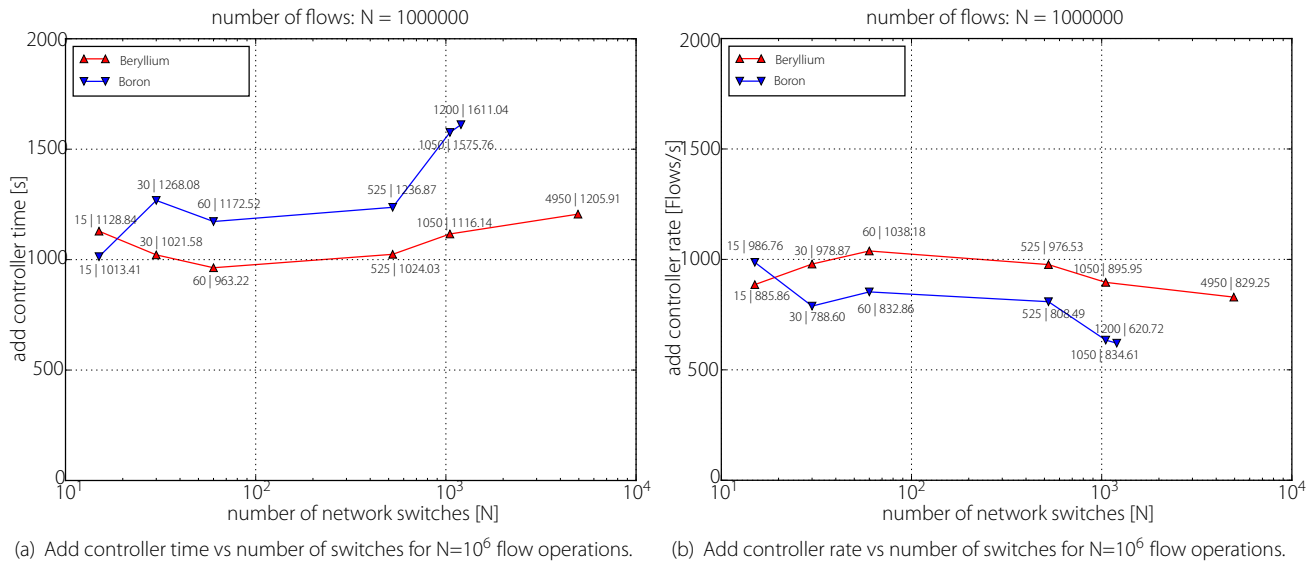


Fig. 19: Flow scalability stress test result. Comparison performance for add controller time/rate vs number of switches for various numbers of flow operations N. Add controller time is forced to -1 when test fails. Add controller time/rate vs number of switches for  $N=10^6$  flow operations.

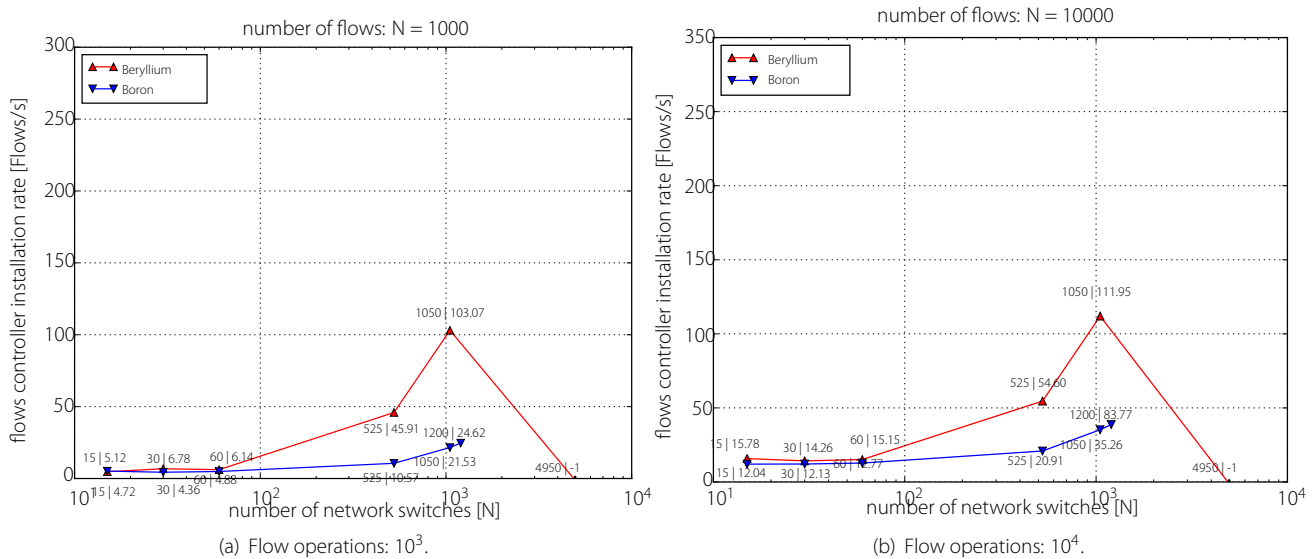
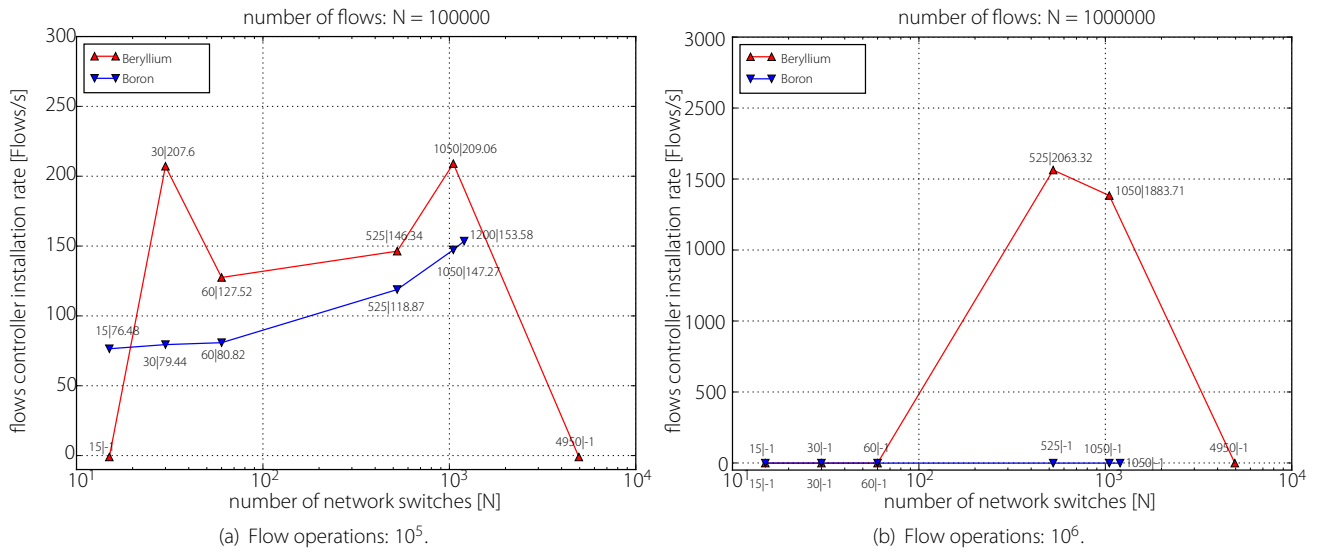


Fig. 20: Flow scalability stress test result. Comparison performance for add controller time/rate vs number of switches for various numbers of flow operations N. Add controller time is forced to -1 when test fails. Add controller time/rate vs number of switches for  $N=10^4$  flow operations.

- [4] "MT-Cbench: A multithreaded version of the Cbench OpenFlow traffic generator." <https://github.com/intracom-telecom-sdn/mtcbench>.
- [5] "NSTAT: NorthBound Flow Emulator." <https://github.com/intracom-telecom-sdn/nstat-nb-emulator>.
- [6] "NSTAT: Performance Stress Tests Report v1.2: "Beryllium Vs Lithium SR3"" [https://raw.githubusercontent.com/wiki/intracom-telecom-sdn/nstat/files/ODL\\_performance\\_report\\_v1.2.pdf](https://raw.githubusercontent.com/wiki/intracom-telecom-sdn/nstat/files/ODL_performance_report_v1.2.pdf).
- [7] "Docker containers." <https://www.docker.com/what-docker>.
- [8] "Docker-hub Intracom-repository." <https://hub.docker.com/u/intracom/>.
- [9] "OpenDaylight Performance: A practical, empirical guide. End-to-End Scenarios for Common Usage in Large Carrier, Enterprise, and Research Networks." [https://www.opendaylight.org/sites/www.opendaylight.org/files/odl\\_wp\\_perftechreport\\_031516a.pdf](https://www.opendaylight.org/sites/www.opendaylight.org/files/odl_wp_perftechreport_031516a.pdf).
- [10] "Mininet. An instant virtual network on your Laptop." <http://mininet.org/>.
- [11] "OpenVSwitch." <http://openvswitch.org/>.



**Fig. 21:** Flow scalability stress test result. Comparison performance for add controller time/rate vs number of switches for various numbers of flow operations N. Add controller time is forced to -1 when test fails. Add controller time/rate vs number of switches for  $N=10^6$  flow operations.

[12] "Generate PACKET\_IN events with ARP payload."  
[https://github.com/intracom-telecom-sdn/multinet#generate-packet\\_in-events-with-arp-payload](https://github.com/intracom-telecom-sdn/multinet#generate-packet_in-events-with-arp-payload).