LSST QoS Policy Test Plan

Table of Contents

LSST QoS Policy....3
LSST Network Traffic Types....3
Topologies....6
Controlled Topology....7
Test Cases....7
Test Case 1 – Assess QoS queues scheduling....7
References....11

LSST QoS Policy

The main objective of this document is to describe a QoS policy that will be utilized in the LSST (Large Synoptic Survey Telescope) project and to specify some test cases that will be executed to validate the switches platforms used in this project, i.e., Corsa DP2400, Brocade MLXe, Brocade SLX and Dell z9100-ON, are able to handle the traffic congestion according to the QoS policy that will be defined in this document.

In addition to the QoS congestion scheduling policy, this test plan will also include test cases for other QoS techniques and mechanisms such as traffic metering and shaping.

LSST Network Traffic Types

Table 1 illustrates the traffic types and the expected level of priority, including the bandwidth that should be guaranteed for the LSST project.

Traffic Types	Priority (Descending order)	La Serena - Santiago (Leased + Lambda)		Santiago to Miami (Leased + Lambda)			Miami to Chicago (Shared + Esnet)		Chicago - Champaign
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
		Bandwidth Available 140 G	Lambda Down BW Avail 40G	Bandwidth Available 300G	Express Down BW Avail 200G	Express + Protect 1 side Down BW Avail 100G	Bandwidth Available 200G	Lambda Down BW Avail 100G	
Science Data Transport [1, 5, 6]	1	35	35	**35**	35	35	35	35	
OCS/TCS/CCS/DMCS [2, 5, 6]	2	20	0	20	20	20	20	20	
Data Backbone Transfer [2, 5, 6]	3	20	0	20	20	20	20	20	
Chilean Community Use [2, 5, 6]	4	20	0	20	20	20	20	20	
Other LSST operations traffic, non-LSST AURA traffic [3, 5, 6]	5	NA	NA	NA	NA	NA	NA	NA	
Bandwidth Reservation for LSST (Lambda/Spectrum + Leased)		25%	12%	25%	38%	75%	38%	75%	0%
NOTES:									
[1] Guaranteed reservation					Shared via Leased				
[2] Expected use, Prioritized/QoS					**Dedicated via Spectrum**				
[3] Carried on AURA circuits, not LSST. Fail-over from AURA to LSST is not baselined.									
[4] Combinations of Cases use numbers from lowest performance case									
[5] All numbers are stated in line speed reservation, efficiency may lower actual delivered BW									
[6] Cases of total network failure" on a link, i.e. all paths connecting two points are not shown. Those cases rely on data buffering and catch-up.									

Table 1 - Qos Traffic Types x Prioritization [1]

In the production network that will transport LSST traffic, there are both 40G and 100G interfaces, and considering the fact that network switches provide QoS prioritization and scheduling queues at the interface level, the QoS policy should be based on the transmission bandwidth of each interface, for example, as it is depicted in Figure 1 and Figure 2. Table 2 represents a particular traffic prioritization and the total bandwidth guaranteed in case of traffic congestion for these traffic types and which queue each traffic type is classified.

Traffic Type	Queue # (8 in total/interface)	40G Interface		100G Interface	
		Gbps	%	Gbps	%
Control & Management plane	7	0.4	1	1	1
Future use (reserved)	5, 6	0.8	2	2	2
Science Data Transport	4	36.8	92	35	35
OCS/TCS/CCS/DMCS	3	0.8	2	22	22
Data Backbone Transfer	2	0.4	1	20	20
Chilean Community Use	1	0.4	1	19	19
Best Effort	0	0.4	1	1	1
		Total Reserved:		Total Reserved:	
		40	100	100	100
		Total Remaining:		Total Remaining:	
		0	0	0	0

Table 2 – QoS Traffic Types x Weights of the Queues

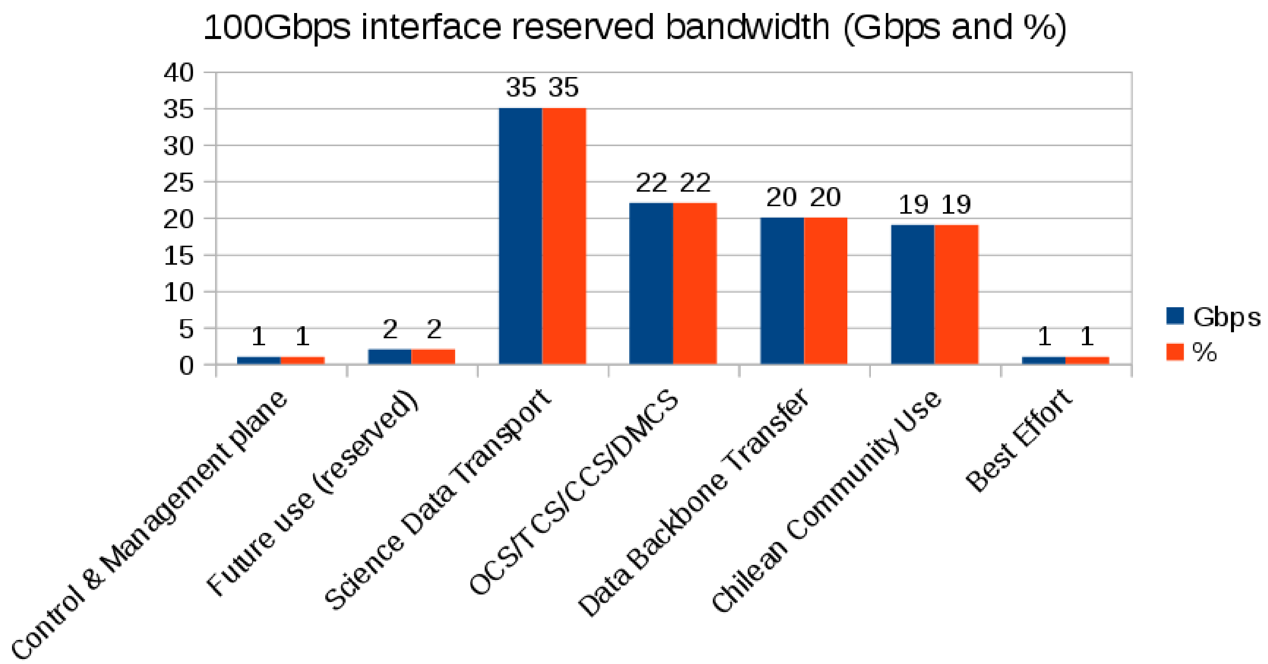


Figure 1 – 100Gbps interface reserved bandwidth (Gbps and %)

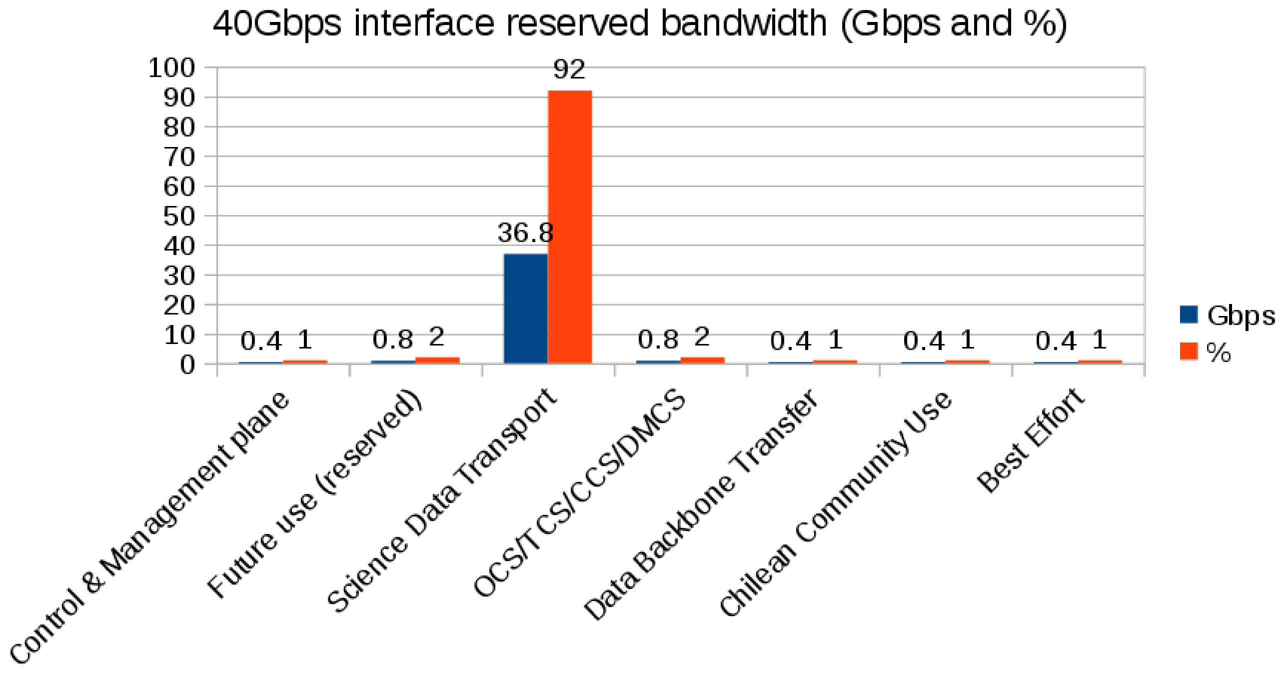


Figure 2 – 40Gbps interface reserved bandwidth (Gbps and %)

Although the proposed LSST production topology is composed of 40Gbps and 100Gbps interfaces, for testing purposes, according to the the laboratory platforms that AmLight has available to start testing, 10Gbps interfaces (Figure 3) will be used to represent 100Gbps interfaces. As a result, the final outcome is expected to be proportional. Nevertheless, 100Gbps interfaces will be tested later on in the production topology.

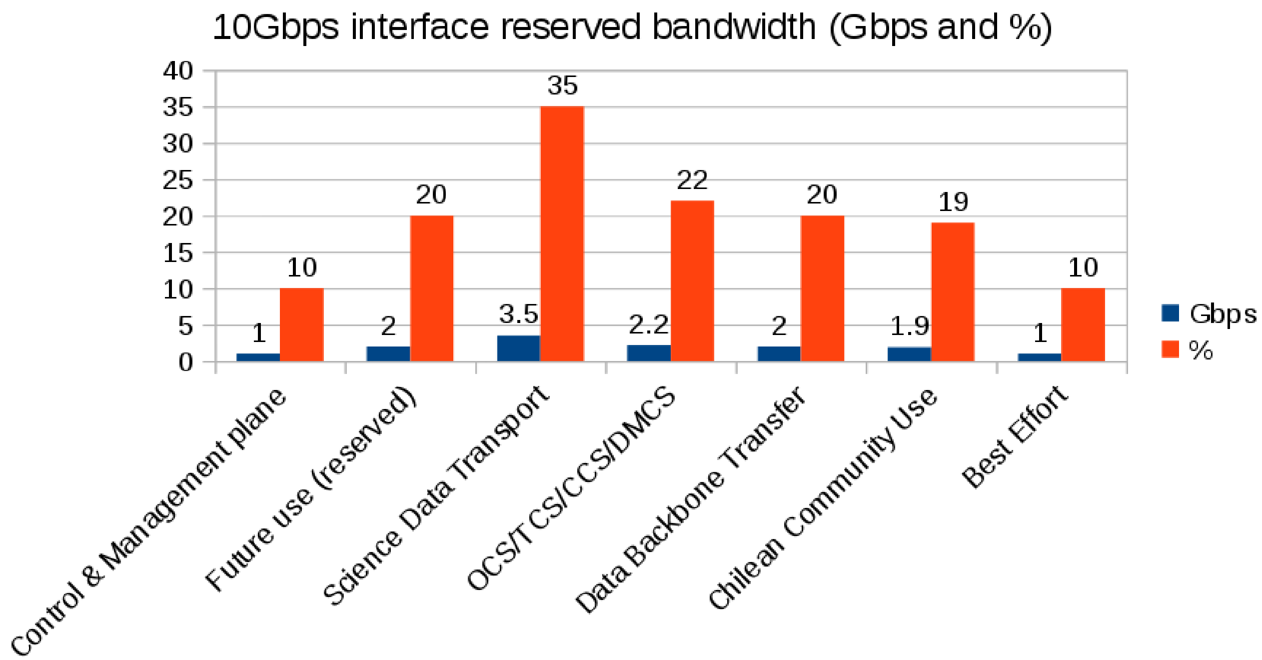


Figure 3 – 10Gbps interface reserved bandwidth (Gbps and %)

Topologies

There are two topologies, controlled (Figure 4) and production, which will be utilized for validating the QoS Policy proposed in this document. The former will be used with the new experimental SDN Corsa DP2400 switch that AmLight has available at AMPATH, one of the reasons to use the controlled topology is to facilitate testing and also to validate Corsa Platform QoS feature set, which is quite promising when compared to other vendors. The controlled topology will act as a switched SDN transport network, with QoS policy applied, to transport traffic from a simulated LSST source at AMPATH to a simulated LSST destination at AMPATH. The latter is the production topology, which will be tested with the same test cases and fully documented, as soon as the testing procedure finishes in the controlled topology. Essentially, the same test cases will be run in the production topology, and AmLight Engineers will update this document accordingly.

In the production topology, there are other SDN switch platforms, e.g., Brocade and Dell, and also 100Gbps interfaces. In addition, in the production topology, the LSST source will be in Chile to a LSST destination in NCSA.

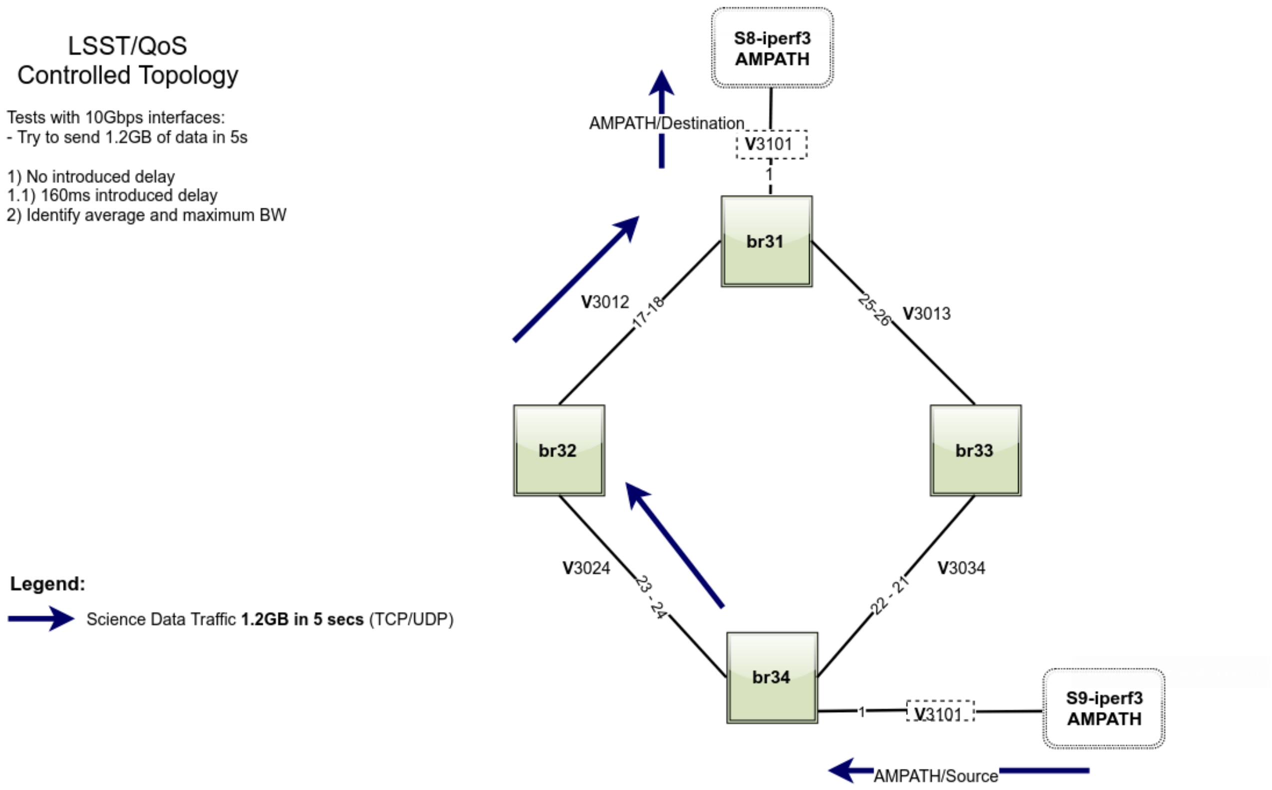


Figure 4 – LSST QoS Controlled Topology

Controlled Topology

Test Cases

Test Case 1 – Assess QoS queues scheduling

The main objective of this test case is to verify that QoS queues scheduling are handling the classified traffic according to the weights of the queues that have been preconfigured. In order to assess this behavior, two types of network traffic will be part of this test case, best effort and science data traffic, which is generated from the server "S8" to the destination "S9". An additional server "S0" running iperf3 will be utilized to test this traffic overload, as you can see in Figure 6. In this case, the best effort queue has 1% of the guaranteed bandwidth, whereas data science have at least 35%.

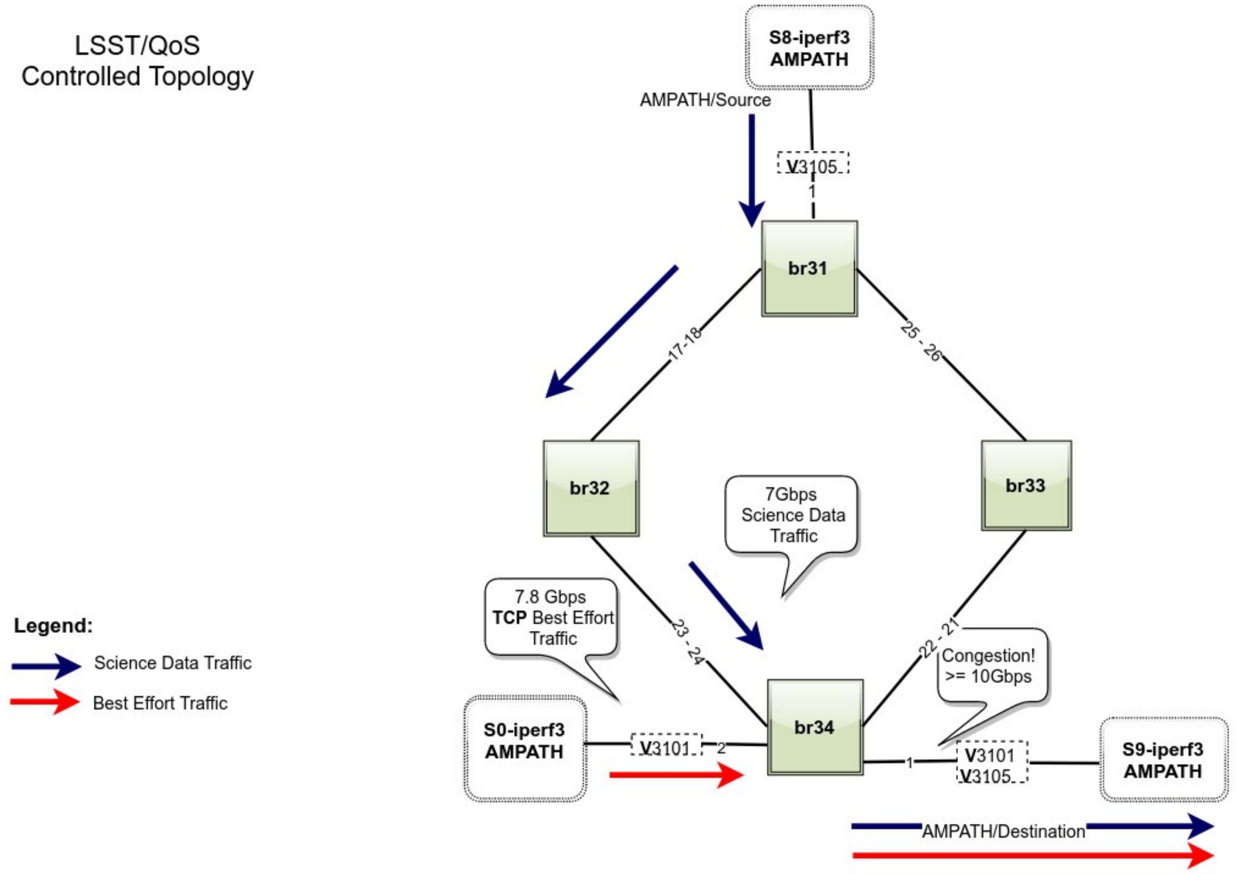


Figure 6 – QoS Controlled Topology of Test Case 2

In summary, the objective of this test cases consist of validating these points:

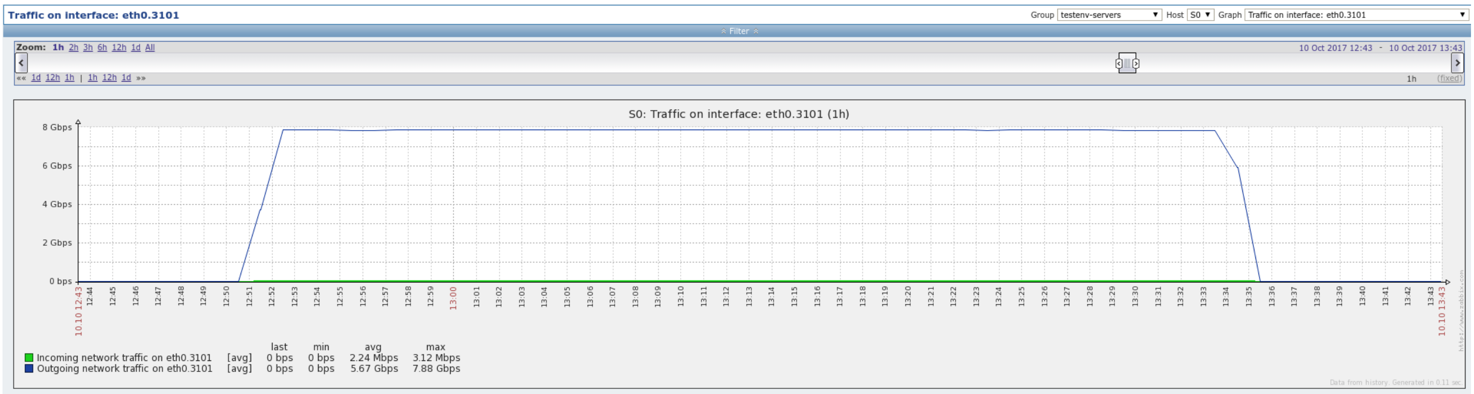
1. Generate both simulated Science Data and Best Effort traffic in order to verify that the congestion queueing scheduling is working as expected.

1.1. Initially, only the Best Effort traffic will be present. Consequently, the Best Effort traffic is allowed to use the remaining bandwidth, which is up to 10 Gbps.

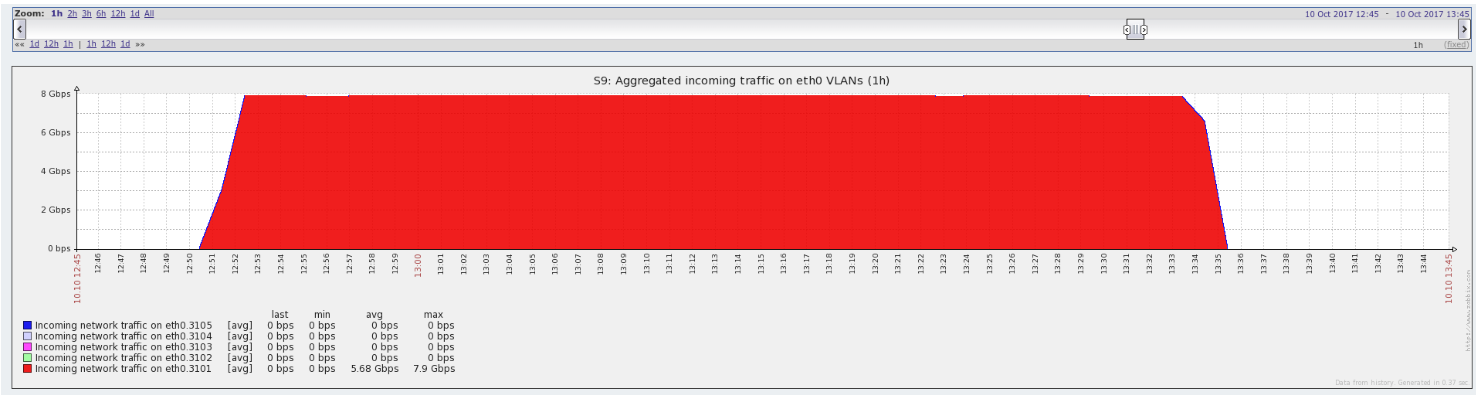
Transmitter Server	**Traffic Type**	**UNI (access) - Transmission Rate (Gbps)**	**NNI (upstream) - Expected Transmission Rate (Gbps)**	**Receiver Server**
S0	Best Effort	7.8	7.8	S9

Result:

Approximately at 12:50 S0 starts the Best Effort traffic type, at roughly 7.8 Gbps:



On S9, the Best Effort is received without any other concurrent traffic at 7.8 Gbps, as expected:

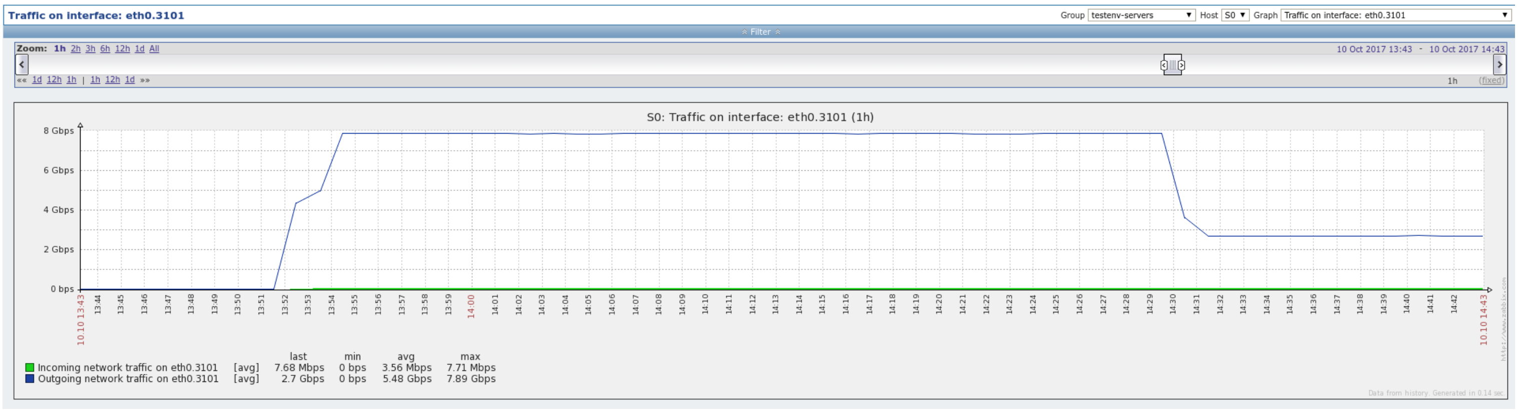


1.2. Overtime, the Science Data traffic will be introduced at 7 Gbps. As a result, since the rate between these two queues is 35:1, Science Data traffic has a higher weight, which allows it to transmit at 7 Gbps and the Best Effort traffic can use the remaining bandwidth (up to 3 Gbps).

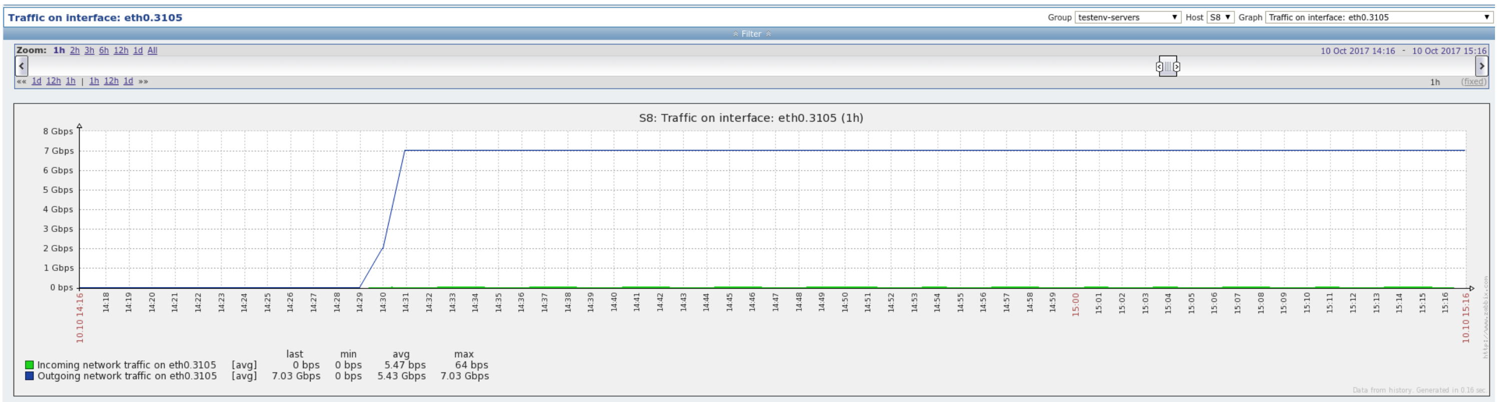
Transmitter Server	**Traffic Type**	**UNI (access) - Transmission Rate (Gbps)**	**NNI (upstream) - Expected Transmission Rate (Gbps)**	**Receiver Server**
S0	Best Effort	7.8	3	S9
S8	Science Data	7	7	S9

Result:

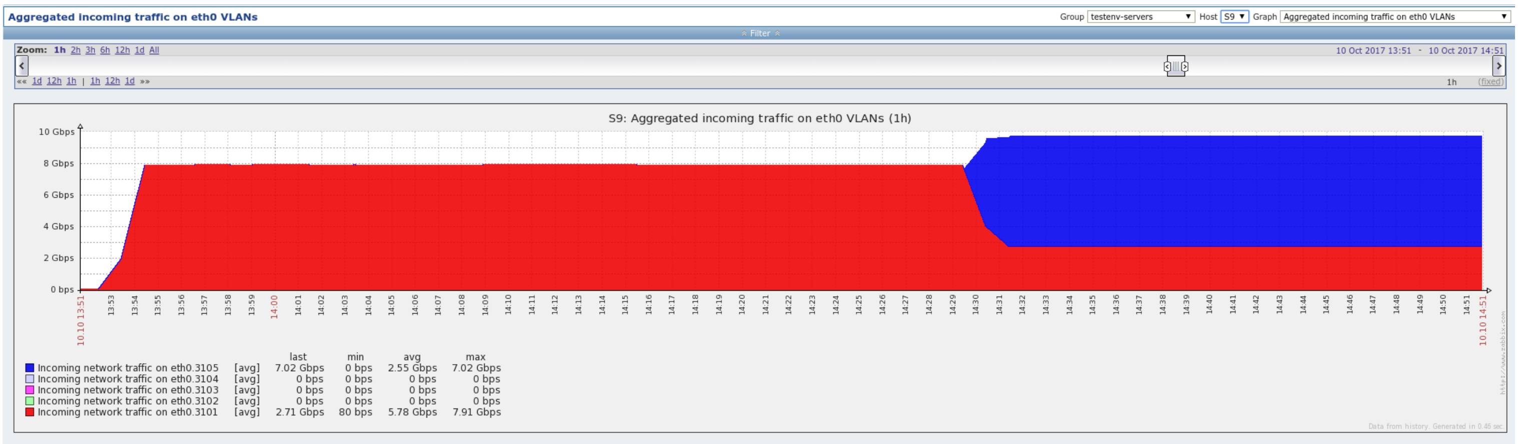
S0 started the 7.8 Gbps Best Effort TCP traffic transmission at approximately 13:54. As it is depicted bellow, at 14:30 the outbound transmission dropped to roughly 3 Gbps, because at 14:30 the simulated Data Science data traffic was introduced on S8.



S8 started transmitting the simulated Data Science traffic at 14:30 at 7 Gbps:



As you can see on the aggregated incoming traffic on S9, the Data Science traffic took over the Best Effort, which used the remaining 3 Gbps bandwidth.



References

[1] LSST Network Bandwidth Allocation. Available at https://confluence.lsstcorp.org/display/DM/LSST+Network+Bandwidth+Allocation, accessed on Sep 22nd.