

Table of Contents

Finding congestion point for monitoring traffic.....	1
Simulated Topology.....	1
Topology Description (topologyGenerator).....	1
Simulation Results.....	2
Congestion point.....	2

Finding congestion point for monitoring traffic

We sweep the amount of monitoring traffic generated by each felix server with 7 simulation. All other parameters are fixed.

A detailed description of each individual simulation is similar to the previous scenario, here we should the aggregated results of the 7 simulations.

Conclusions

1. When the monitoring traffic increases, plots indicate that the congestion point is the hlt_core routers. The maximum monitoring traffic that this network can absorb is 0.66 Gbps per felix server. Although latency starts increasing rapidly from 0.6 Gbps onwards.
2. Latency for the monitoring traffic stays below 0.001 us when there is no congestion.
3. The maximum queue length in the HLT routers reached 400KB (per output) at the congestion point (0.66Gbps) and at 100KB for (0.6Gbps)
4. the data-flow latency was not affected by the increase in the monitoring traffic because the only shared link is with <50% usage

Simulated Topology

Based on this meeting:

<https://docs.google.com/presentation/d/1N5GnG82JvsASUJ6sl-gmuMBRjMWG6OWtybIjOMikb-M/edit#slide=id.g1>

We simulate the full path of the monitoring traffic all the way to the monitoring servers in the HLT farm.

We sweep the amount of monitoring traffic generated by each felix server from 0.4Gbps to 0.8Gbps (the congestion point is ~0.633Gbps) . All other parameters are fixed (32 Gbps of data-flow traffic)

Monitoring flows (NFelix -> 1 MonServer)

Topology Description (topologyGenerator)

NUMBER_OF_FELIX_SERVERS = 13 # this generates 1:1 connections with sw_rod, so

NUMBER_OF_FELIX_SERVERS=numberOfSWRODServers

NUMBER_OF_MONITORING_SERVERS = 5

LINK_BW_40G_BITS_S = 40 * G # 40 Gbps

LINK_BW_1G_BITS_S = 1 * G # 1 Gbps

FELIX_FLOW_PRIORITY = 0

FELIX_GBT_ELINKS = 10 # #GBT e-links in each felix server. There will be one flow created per e-link (because there will be 1 thread, one connection per e-link)

felix data-flow distributions (one per GBT)

FELIX_GBT_PERIOD_sec = ExponentialDistribution.new 1.0 / (100*K) # distribution period in seconds

FELIX_GBT_SIZE_bytes = NormalDistribution.new 4.0*K, 1.0*K # (in bytes)

FELIX_GBT_BUFFER_bytes = 1*M # (in bytes)

FELIX_GBT_TIME_OUT_sec = 2 # (in seconds)

FELIX_GBT_OUT_SIZE_bytes = TCP_MTU_bytes # (in bytes)

FELIX_GENERATION_PERIOD = FelixDistribution.new FELIX_GBT_PERIOD_sec,

FelixDistribution::FELIX_MODE_HIGH_THROUGHOUT,

FELIX_GBT_SIZE_bytes,

```
FELIX_GBT_BUFFER_bytes,  
FELIX_GBT_TIME_OUT_sec,  
FELIX_GBT_OUT_SIZE_bytes  
FELIX_GENERATION_SIZE = ConstantDistribution.new TCP_MTU_bytes*8 #distribution size in bits
```

```
# monitoring flows (one per GBT)  
FELIX_MONITORING_PRIORITY = 0  
MONITORING_SIZE_bits = (TCP_MTU_bytes - 300)*8  
TOTAL_MONITORING_PER_SERVER_bits = 0.3 * G --> this is the value we are sweeping  
MONITORING_GENERATION_PERIOD = ExponentialDistribution.new 1.0 /  
(TOTAL_MONITORING_PER_SERVER_bits / (MONITORING_SIZE_bits * FELIX_GBT_ELINKS))  
MONITORING_GENERATION_SIZE = NormalDistribution.new MONITORING_SIZE_bits, 300*8  
#distribution size in bits
```

Simulation Results

A detailed description of each individual simulation is similar to the previous scenario, here we should the aggregated results of the 7 simulations. I.e: each point plotted in these figures represent a complete simulation execution.

In the plots, the aggregation is usually done by averaging all the values from the same simulation. For example:

1. **The average** of all values in the simulation is represented with a point in the plot (for example in the "max queue" plot, each point it is the average of all maximum queue lengths).
2. **The std deviation** is plotted with a light-blue shade
3. **The maximum** values are plotted with red bars.

The plot shows scenarios with and without congestion in order to find the maximum traffic that the network can absorb. For the simulations where congestion occurs, the results are not realistic as buffers are configured with infinite capacity.

Congestion point

When the monitoring traffic increases, plots indicate that the congestion point is the hlt_core routers. The maximum monitoring traffic that this network can absorb is 0.66*G per felix server. Although latency starts increasing rapidly from 0.6 Gbps onwards.

mean latency

The plot shows that the latency for the monitoring traffic stays below 0.001 us when there is no congestion. The maximum monitoring traffic that this network can absorb is 0.66*G per felix server. Although latency starts increasing rapidly from 0.6 - 0.63 Gbps onwards.

hlt cores

When the monitoring traffic increases, this plot indicate that the congestion point is the hlt_core routers. The maxium queue lenght in the HLT routers reached 400KB (per outport) at the congestion point (0.66Gbps) and at 100KB for (0.6Gbps)

mean latency

The following plot shows that the data-flow latency was not affected by the increase in the monitoring traffic.

This is because the data-flow and monitoring-flow share only the Felix NIC output queue and this link is used only <50% (the link 2x40Gbps, and there is 32 Gbps of data-flow traffic and max 0.8 Gbps of monitoring)

-- MatiasAlejandroBonaventura - 2016-12-12

This topic: Main > PhaseISimulation_LArSliceMonitoringSweep

Topic revision: r3 - 2016-12-13 - MatiasAlejandroBonaventura



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback