

Project Proposal

Implement the query-based network monitoring application
(from Section 4.5, Exercise 2.)

Chenye Yang cy2540

Bingzhuo Wang bw2632

Zhuoyue Xing zx2269

Data / Input

According to the requirement, we need to develop a network flow workload generator as a source operator.

1. the rate of incoming tuples summarizing an individual flow
2. the number of IP addresses and protocols involved in the flows
3. the distribution of the network usage over the different kinds of protocols and IP addresses

Data generated and sent in real-time.

Techniques / Applications

Problems:

Monitor the network resources and answer some statistical questions.

1. List protocols that are consuming more than H percent of the total external bandwidth over the last T time units.
2. List the top- k most resource intensive protocols over the last T time units.
3. List all protocols that are consuming more than X times the standard deviation of the average traffic consumption of all protocols over the last T time units.
4. ...

Techniques:

GUI application and Spark program

Results

A GUI-based application will be our demo.

Record a video of the GUI.

Novelty

Develop a system which:

1. Process network data and visualize network resources usage;
2. Be deployed to server and accessed remotely;
3. Use watermark to ensure fault tolerance.