**Graph based SIEM Analysis**

User Guide

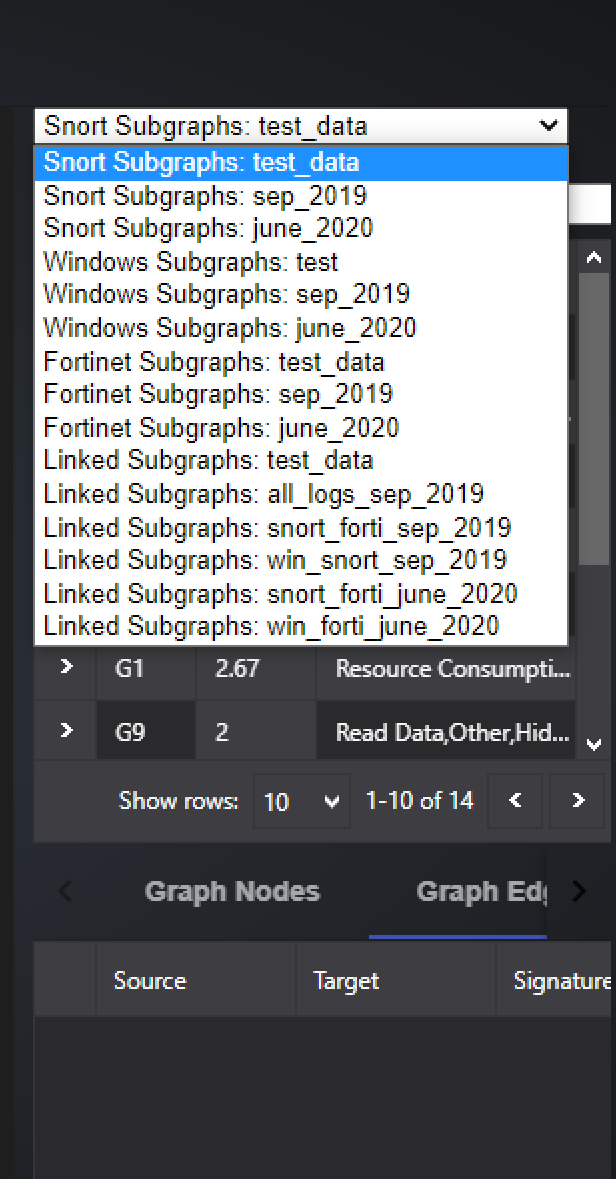
**Overall intention and working of the model for Graph based SIEM Analysis**

The model as of now takes as input SIEM alerts for 3 log sources – Snort, Fortinet and Windows. It ingests the alerts for a configurable period of time (set at 1 month for now), and generates graphs based on them. These graphs represents collections of similar activities seen across the log types. We consider activities to be similar, if the are similar in terms of the signature which led to generation of alerts, as well as similarity in terms of the intensity of the events, duration of the events as well as whether the events target specific ports or were generated from similar ports or not.

The intention of generating these graphs is to identify patterns of similar activity from very voluminous SIEM alerts. This can be helpful, as SIEM alerts can be very voluminous with thousands of alerts generated on a daily basis. In the data we have, we see Snort logs have 4000 to 16000 alerts generated on a daily basis. Due to these large volumes, it is difficult for analysts to identify patterns through manual analysis at the alerts level. By consolidating collections of similar activities within and across log types, we provide an easier tool for inference to the analysts. Further, the graphs are also sorted by a Severity Score calculated by the model. This Severity score gives the analyst an indication of the maliciousness of the events seen in each graph.

**Q 1 .How to view list of graphs generated for a particular time period and log source?**

Currently the demo shows graphs generated for September 2019 and June 2020. Our log sources are SIEM alerts for Snort, Fortinet and Windowss.

You can view the list of graphs for a particular log source and month by the dropdown:

For example, to view list of graphs generated for Snort alerts, for September 2019, select Snort Subgraphs: sep\_2019.

**What are linked Subgraphs?**

Linked graphs are graphs generated by connecting events across log types - Snort, Fortinet and Windows.

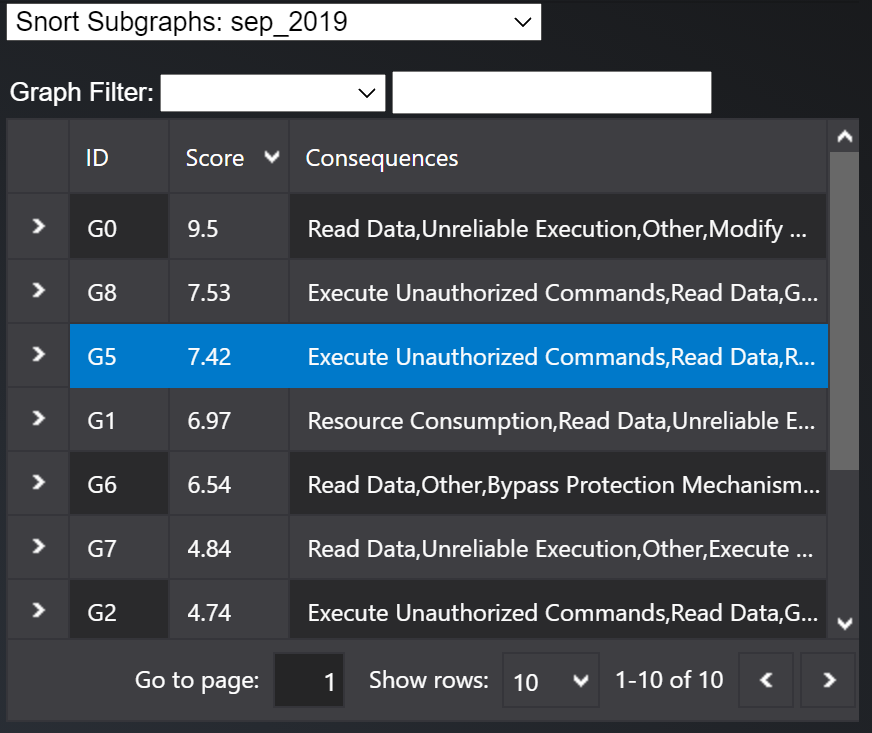
Hence, if graphs generated in Snort and Fortinet have IP addresses in common, then linked graphs between Snort and Fortinet would be generated by connecting these individual graphs, re-calculating the graph severity and consequences.

Linked Subgraphs all\_logs\_sep\_2019 contains those graphs which have nodes (IP addresses) found in alerts across all 3 log types - Snort, Fortinet and Windows.

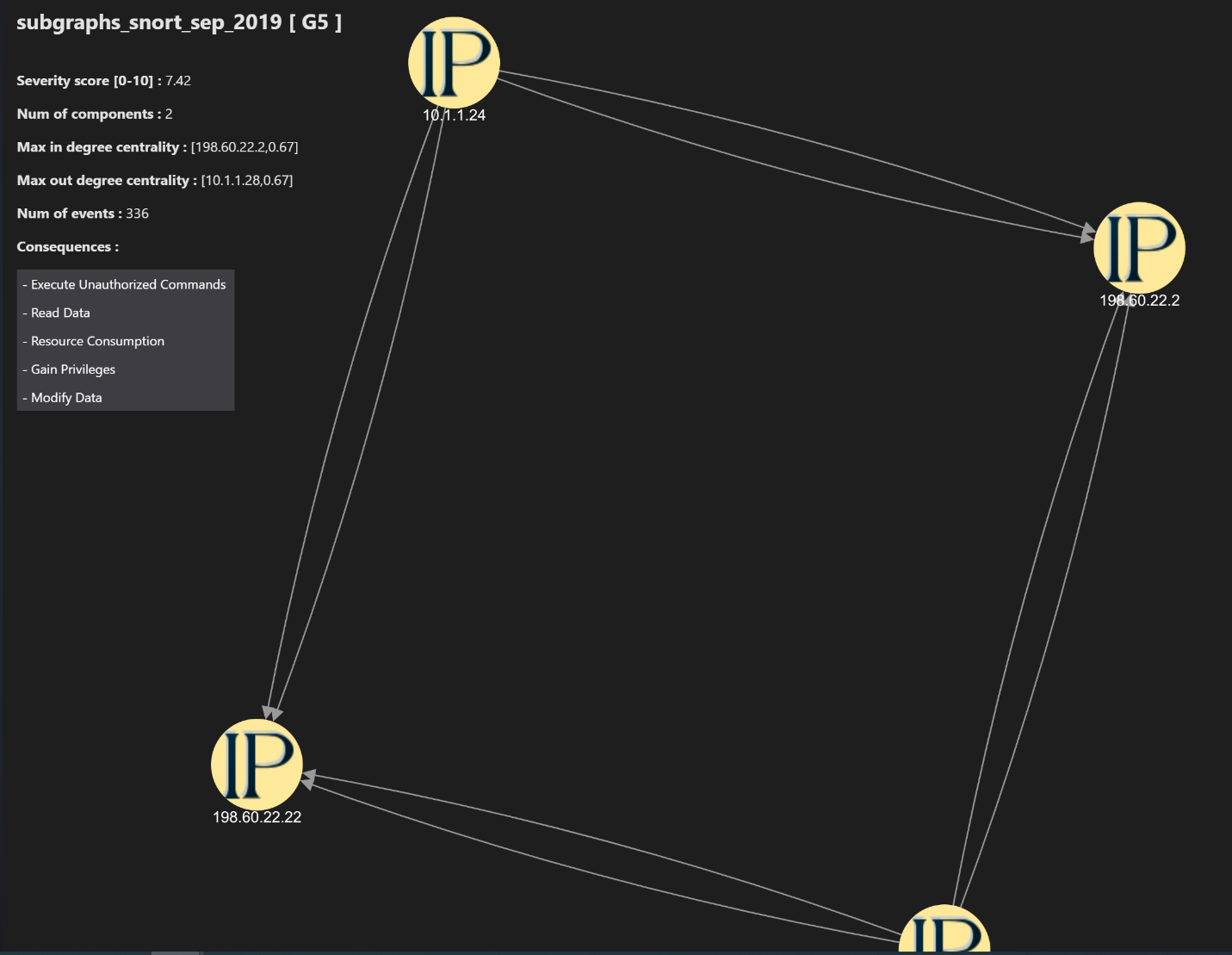
Linked Subgraphs Snort\_forti\_sep\_2019 contains those graphs which have nodes (IP addresses) found in alerts across all Snort and Fortinet.

Linked Subgraphs win\_snort\_sep\_2019 contains those graphs which have nodes (IP addresses) found in alerts across all Snort and Windows.

**Q2. How to view a particular graph?**

Once the user selects a log source and time period from the dropdown as shown in Q1, the list of subgraphs are shown below. For example, if the user selects Snort Subgraphs : sep\_2019, the list of graphs generated for Snort alerts for September 2019 would be shown as below:

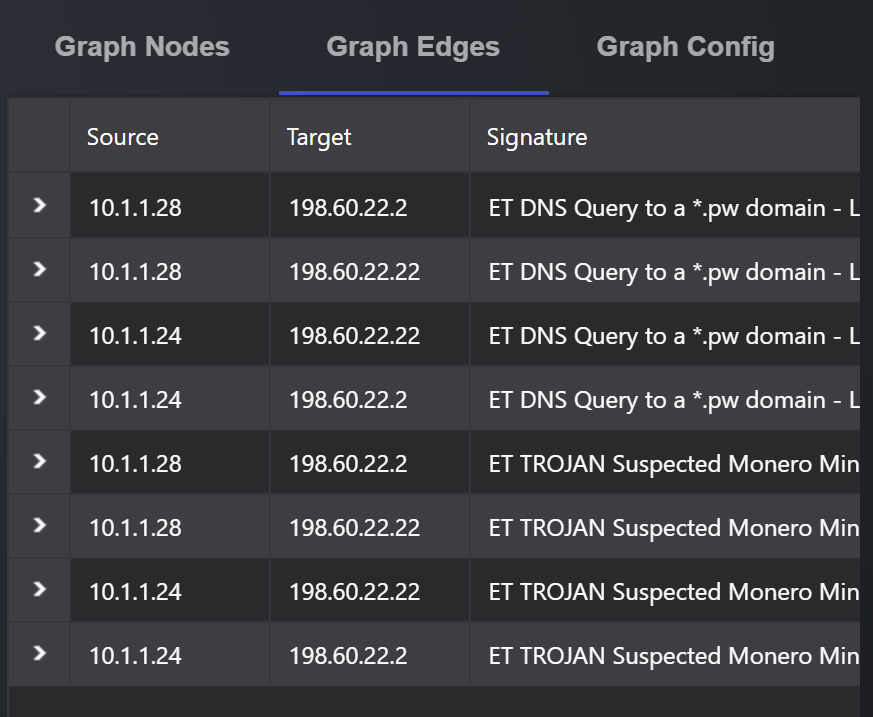
The list provides a graph ID (eg. G5), Severity score of the graph (eg. 7.42) and list of possible consequences identified based on the types of signatures seen in the graph (eg. Execute Unauthorized commands, Read Data etc).

The left panel of the UI shows the graph and related information:

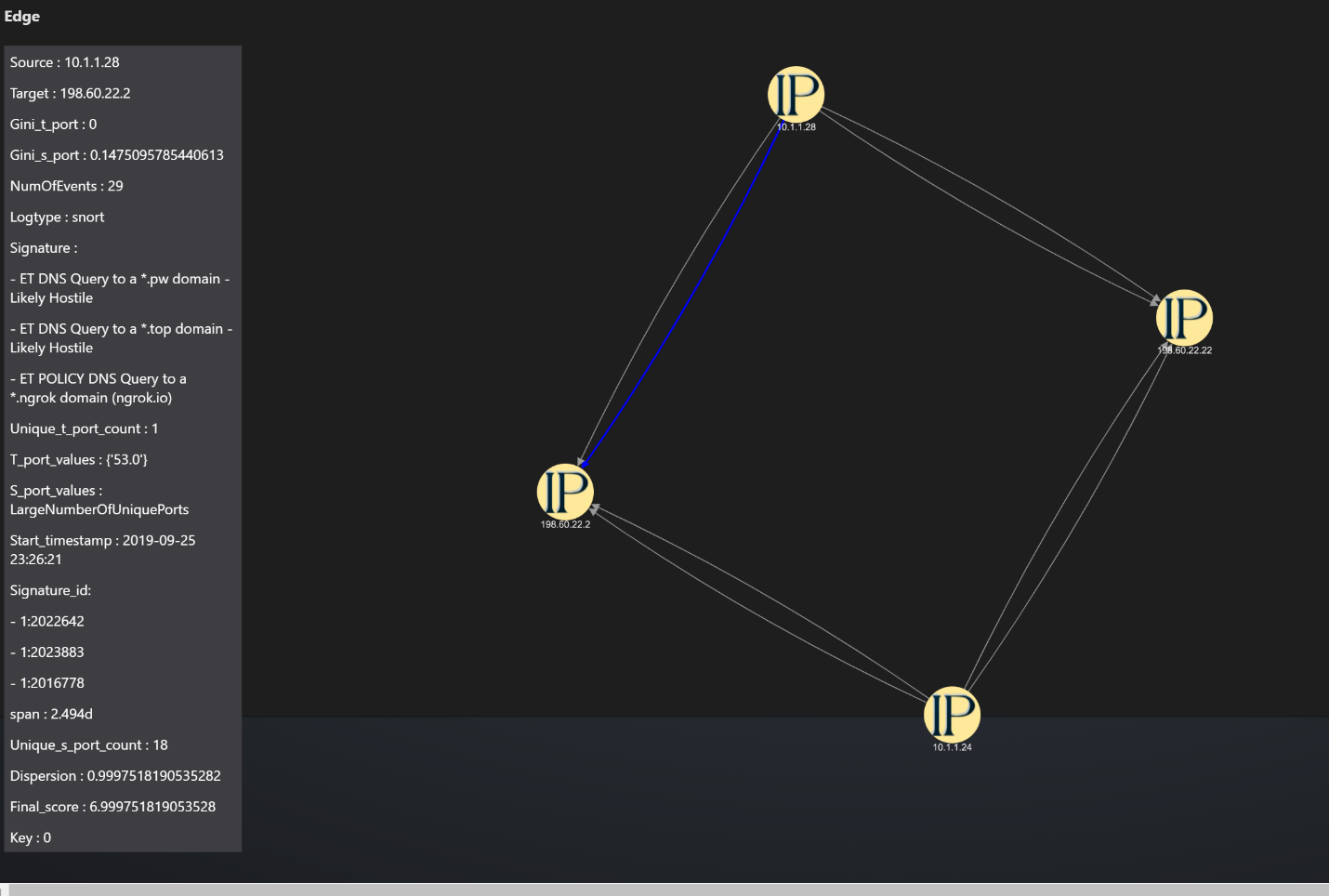
The top left portion of the left panel gives high level information for each graph:

* Severity score : Severity score of the graph calculated by the model.
* Num of events : Total number of SIEM alerts consolidated by the model to build the graph. For example, in the screen shot above, graph G5 consolidates information seen in 336 individual alerts in the SIEM data. If the analyst did not have access to graph G5, they would have to infer this information by analyzing 336 alerts themselves.
* Num of components : Each graph represents a certain type of activity seen across a log type for 1 month. For example, in G5 the graph consolidates suspicious DNS queries seen across Snort logs. Such activities can be seen across 1 collection of IP addresses or multiple collections. The Num of Components tells the analyst the spread of such activity – how many collections of IP addresses are experiencing this activity?
* Max in degree centrality – Which node (IP address / Computer Name) has the maximum number of incoming edges? If the value is high ( near 1 or higher than 1), that implies most of the edges are targeted to a specific node. If the value is less, it implies the targets are a range of IP values.
* Max out degree centrality : Similar to Max in degree centrality, tells if the source for most edges is 1 or 2 IP values, or its spread across various values.
* Consequences : Possible types of consequences which may result based on the signature and other activity patterns seen in the graph.

**Q3 How to view graph details?**

Details on the graph edges can be viewed in the right side panel under Graph Edges:

Selecting a particular edge from the Graph Edges section highlights the edge in the graph visualization as well. Each edge has the following information:



Signature : Which signatures have been seen in the alerts consolidated in the edge.

NumOfEvents : How many alerts from the SIEM logs have been consolidated to form the edge.

Start\_timestamp : Earliest timestamp at which the alerts were seen.

Dispersion : Intensity of alerts – were the alerts generated in a burst of activity, or were periodic in nature. Value closer to 1 indicates high intensity.

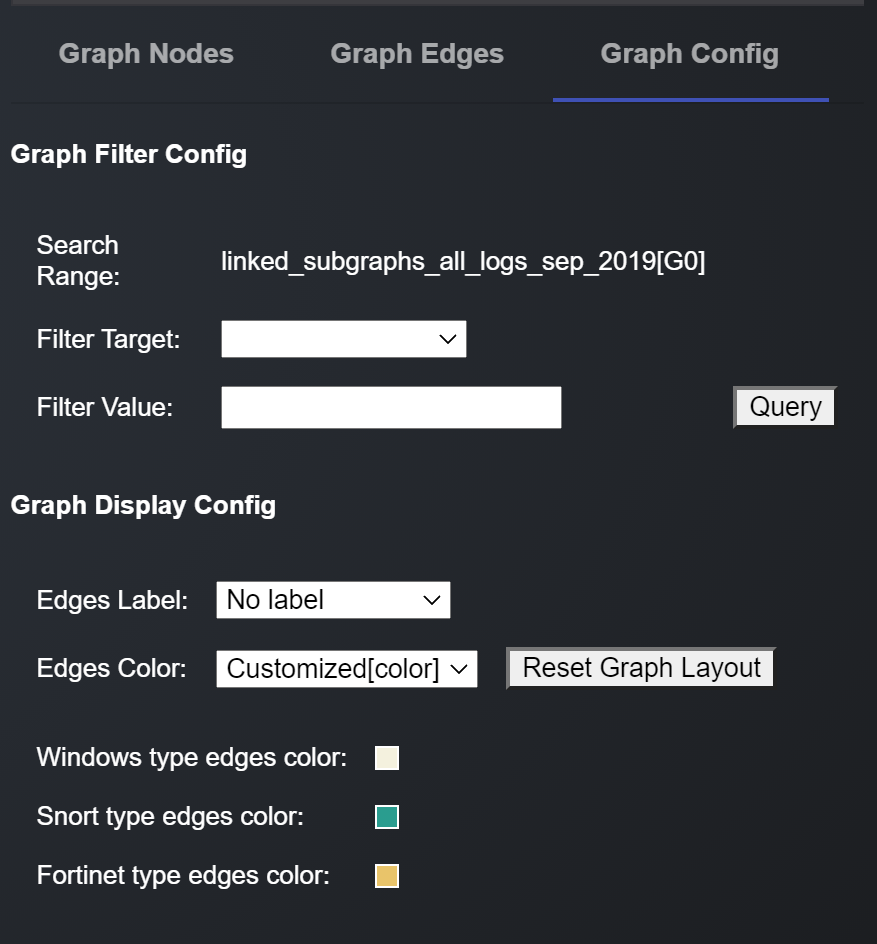
Span : Duration for which alerts were seen through the month (for ex. 2.49d - ~2.5 days).

S\_port\_values and T\_port\_values : Source and Target ports seen in the alerts.

Gini\_s\_port and Gini\_t\_port : Wherever large number of ports are seen, were the ports randomly chosen, or concentrated in a few port values. Value closer to 0 indicates the ports were randomly chosen.

**Q4 How to configure graph visualization?**

Graph visualization can be configured through the Graph Config tab in right side panel:



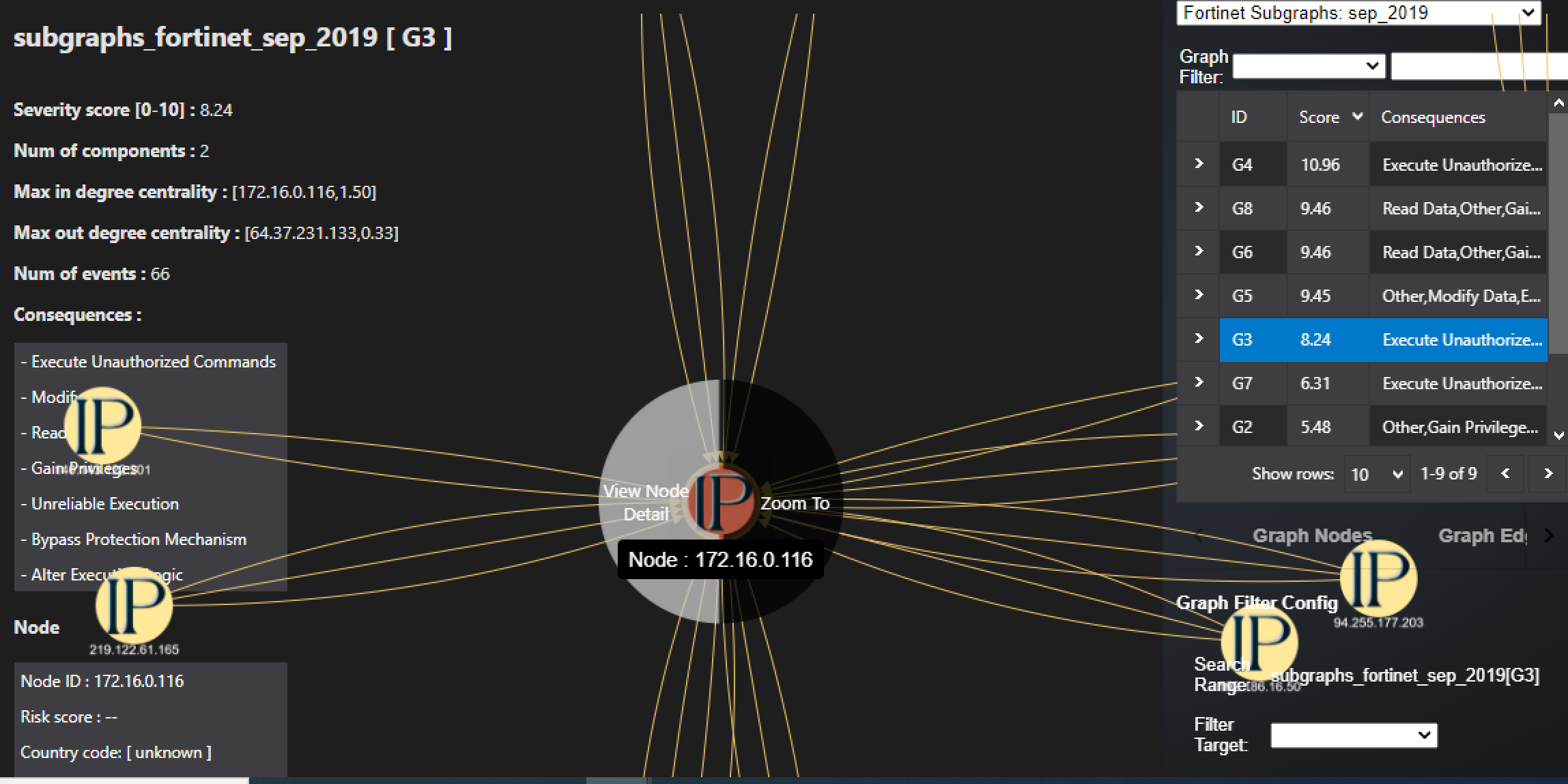
User can choose to see Signature / Log Type / Port Values etc as the edge labels by selecting from the dropdown against Edges Label in the **Graph Display Config**.

Similarly, the edge colors can represent the log types which have resulted in the creation of the edge. This functionality is useful when viewing linked graphs.

The user can also filter edges to view in a graph through the **Graph Filter Config** area.

Q5 How to view all graphs associated with a node (IP address / ComputerName)

Certain nodes are seen across multiple graphs. To view all graphs in Fortinet associated with an IP address (eg. 172.16.0.116), select the node and press left button for some time, and the click on View Node Detail :



This leads to the Node Detail Information page.

This page lists all graphs in Fortinet where this IP is seen in the right side panel.

It also shows all edges and IP address associated with 172.16.0.116 in the graph area, and provides details of the edges in the bottom panel:

