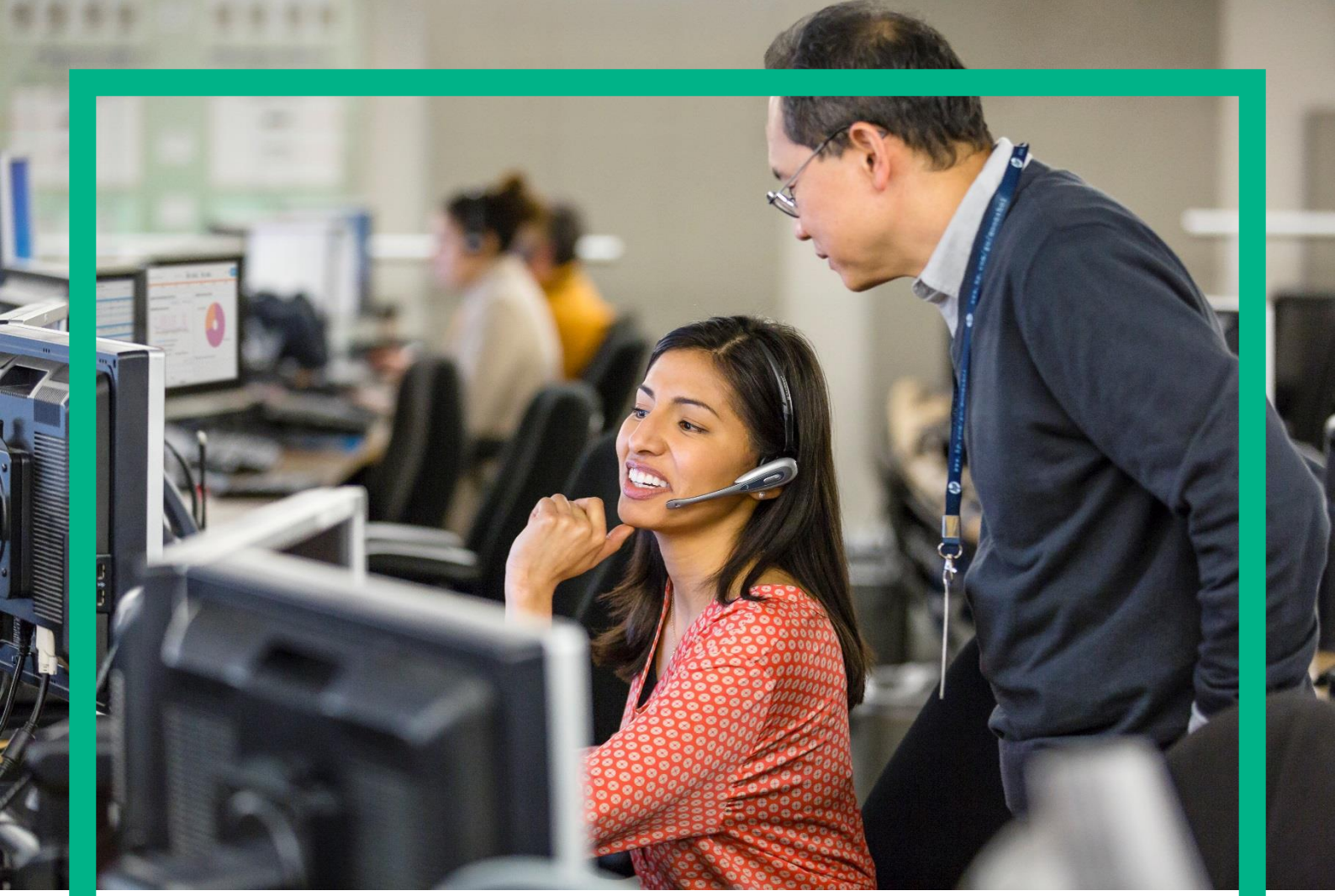


HPE OSS Customer Experience Assurance V5.3

Base Platform
Data Sources and Functions

Hewlett Packard Enterprise

Notices

Legal notice

Printed in the US

Trademarks

Contents

Notices 1

Chapter 1 Introduction 3
1.1 Overview 3

Chapter 2 Data sources 4
2.1 How to use the field matrix 4
2.2 Data sources for aggregated info 5
2.2.1 PS/Zhilabs 6
2.2.1.1 User-Plane 7
2.2.1.2 Control-Plane 22
2.3 Data sources for raw XDRs 25
2.3.1 PS/Zhilabs 26
2.3.1.1 User-Plane 26
2.3.1.2 Control-Plane 39
2.3.1.3 CRM 41

Chapter 3 Functions 42
3.1 Data Range and Data Source 42
3.1.1 Configuring Data Range and Temporal Aggregation 43
3.2 Selecting a data source 44
3.3 Sampling 45
3.4 Filtering by a subscriber or subscribers 46
3.5 Filters 46
3.6 Functions 47
3.6.1 chart 47
3.6.2 cdf 55
3.6.3 pdf 57
3.6.4 download 58
3.6.5 filter 59
3.6.6 get-sessions 60
3.6.7 head 62
3.6.8 heatmap 63
3.6.9 kqi-chart 65
3.6.10 kqi-heatmap 68
3.6.11 ladder 69
3.6.12 ladder-tree 70
3.6.13 map 71
3.6.14 multi-source 72
3.6.15 network 73
3.6.16 pivot-table 76
3.6.17 unique 78
3.6.18 where 79
3.6.19 reduce 80
3.7 Useful tips 93
3.7.1 Reduce 93

Chapter 1
Introduction

1.1 Overview

This document is the data source and functions for Customer Experience Assurance (CEA). The following subsections provide a full description of the available data sources and functions throughout the Query Console.

Chapter 2
Data sources

The data source is each of the different kind of data the CEA is feed of. Data source selection is done by typing:

sourcetype=<datastore_name>

CEA includes an on-line help so that when you write the equal symbol the list of available data sources will be displayed making the selection easier by just clicking on the one you are interested in.

The list of available data sources depends on the acquired license and the integrated data sources.

The next sections depict the different type of data sources available in the system: **aggregated info** and **raw XDRs.**

2.1 How to use the field matrix

The next sections contain a matrix in order to classify every field of each table creating a quick access to see that information.

This section explains the meaning of each column and how to use the following tables:

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field

FIELDS: Name of the field.

Decoration: It marks if the field has been decorated by CEA on mediation layer throughout the decoration files. This process has been created to improve the meaning of some fields making them more human-readable. For example, some IP addresses will be translated into their network element names.

From Network: These fields are directly extracted from the network in the probe servers. No decoration has been applied to them.

Subscriber identifier: The fields under this category are always related to subscriber identifiers so they are the best way to identify a subscribers unambiguously in the datasources for raw XDRs.

Indicator: These fields are used in order to compute a KQI. Some of them could be useful for some queries but the grey ones are designated for internal use exclusively.

Timestamp: These fields contains the exact time when the event has happened in YYYY-MM-DDThh:mm:ss format.

Location field: The fields under this category could be used in order to locate the target register.

CRM field: These fields are always related to customer classification such us data or price plans.

Duration field: These fields are used for store the time duration. The units could be different depending on the target field (milliseconds, seconds, minutes…).

2.2 Data sources for aggregated info

In these data sources the information has been aggregated throughout some metrics for specific dimensions. Therefore these data sources have been optimized to show specifically the preconfigured reports (they are accessible through the “reports” menu at the GUI).

There is a nomenclature agreement to determine which data sources are preconfigured: its name must start with “summary.”

Information is available in data stores in near real time (roughly 15 minutes since data is generated by the network). Default retention periods in the data stores are: 90 days for the lowest aggregation period (15 minutes), 1 year for data aggregated hourly and 2 years for data aggregated daily. Availability periods and retention periods are configurable and should be decided in the Pre-sales Phase, as it has a huge impact in the sizing of the solution.

Note: The GUI will change automatically between the different summarization tables depending on the selected period of time and timeslice. For example, for a 5 hours period with 15 min timeslice it will look inside 900 secs aggregation table but changing the timeslice to 1 hour automatically it will look inside the 3600 secs one.

There are generic fields in most of data sources:

- samples: refers to internal counters used in many different indicator computation.
- rsp: refers to internal response times used in many different indicator computation.

Current data sources list:

Browsing

- summary.ps_browsing-device-900-timeline
- summary.ps_browsing-location-900-timeline
- summary.ps_browsing-network-900-timeline
- summary.ps_browsing-url-900-timeline

DNS

- summary.ps_dns-device-900-timeline
- summary.ps_dns-location-900-timeline
- summary.ps_dns-network-900-timeline

File-sharing

- summary.ps_file_sharing-device-900-timeline
- summary.ps_file_sharing-location-900-timeline
- summary.ps_file_sharing-network-900-timeline

Session

- summary.ps_session-apn-900-timeline
- summary.ps_session-device-900-timeline
- summary.ps_session-location-900-timeline
- summary.ps_session-network-900-timeline
- summary.ps_session-dataplan-900-timeline

Signaling

- summary.ps_signaling-apn-900-timeline
- summary.ps_signaling-device-900-timeline
- summary.ps_signaling-location-900-timeline
- summary.ps_signaling-network-900-timeline

Streaming

- summary.ps_streaming-device-900-timeline
- summary.ps_streaming-network-900-timeline

TCP

- summary.ps_tcp-device-900-timeline
- summary.ps_tcp-location-900-timeline
- summary.ps_tcp-network-900-timeline

The name of this data sources is composed using the following structure:

summary.<family>-<dimension_type>-timeline

The following sections depict the available tables of this data source and list all the fields for each table.

Check the appendix of this document for a further explanation of each field.

2.2.1 PS/Zhilabs

The following data sources store data that has been extracted from Zhilabs App level probes.

2.2.1.1 User-Plane

These data sources contain user-plane information related to data and app traffic.

2.2.1.1.1 Session

summary.ps_session-apn-900-timeline

It contains information related to sessions such as volume or throughput for each APN. The information is aggregated by APN and rat-type.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
bearer.apn		X						
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.activity-volume-down				X				
net.activity-volume-down-samples				X				
net.activity-volume-up				X				
net.activity-volume-up-samples				X				
net.peak-throughputdown				X				
net.peak-throughputdown-samples				X				
net.peak-throughputup				X				
net.peak-throughputup-samples				X				
timestamp					X			

summary.ps_session-device-900-timeline

It contains information related to sessions such as volume or throughput for devices (brand, model and type). The information is aggregated by device brand, device model, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				

bearer.volumeup				X				
bearer.volumeup-samples				X				
device.brand	X							
device.model	X							
device.type	X							
imei.tac		X						
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.activity-volume-down				X				
net.activity-volume-down-samples				X				
net.activity-volume-up				X				
net.activity-volume-up-samples				X				
net.peak-throughputdown				X				
net.peak-throughputdown-samples				X				
net.peak-throughputup				X				
net.peak-throughputup-samples				X				
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

summary.ps_session-location-900-timeline

It contains information related to sessions such as volume or throughput for each location such as cells, city or region. The information is aggregated by ENODB, cell, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.cell	X							
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
cluster	X							
demographics.city	X							
demographics.region	X							
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
location.ci						X		
location.lac						X		
location.mcc						X		
location.mnc						X		
location.sac						X		

net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.activity-volume-down				X				
net.activity-volume-down-samples				X				
net.activity-volume-up				X				
net.activity-volume-up-samples				X				
net.peak-throughputdown				X				
net.peak-throughputdown-samples				X				
net.peak-throughputup				X				
net.peak-throughputup-samples				X				
node.enodeb-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
site-id	X							
timestamp					X			

summary.ps_session-network-900-timeline

It contains information related to sessions, aggregated by PGW, MME, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.activity-volume-down				X				
net.activity-volume-down-samples				X				
net.activity-volume-up				X				
net.activity-volume-up-samples				X				
net.peak-throughputdown				X				
net.peak-throughputdown-samples				X				
net.peak-throughputup				X				
net.peak-throughputup-samples				X				
node.mme-name	X							
node.pgw-name	X							

operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

summary.ps_session-dataplan-900-timeline

It contains information related to sessions, aggregated by dataplan, application and category.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
crm-index							X	
primary_offering							X	
sub-type							X	
timestamp					X			

2.2.1.1.2 DNS

summary.ps_dns-device-900-timeline

It contains information related to DNS operations, aggregated by device brand, device model, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
code-samples				X				
device.brand	X							
device.model	X							
device.type	X							
error-code-samples				X				
imei.tac		X						
net.latency				X				
net.latency-samples				X				
operator-country	X							
operator-id		X						

operator-name	X							
operator-tadig-code	X							
timestamp					X			

summary.ps_dns-location-900-timeline

It contains information related to DNS operations, aggregated by ENODB, cell, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.cell	X							
bearer.rat-type-str	X							
cluster	X							
code-samples				X				
demographics.city	X							
demographics.region	X							
error-code-samples				X				
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
location.ci						X		
location.lac						X		
location.mcc						X		
location.mnc						X		
location.sac						X		
net.latency				X				
net.latency-samples				X				
node.enodeb-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
site-id	X							
timestamp					X			

summary.ps_dns-network-900-timeline

It contains information related to DNS operations, aggregated by PGW, MME, rat-type an operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
code-samples				X				
error-code-samples				X				
net.latency				X				
net.latency-samples				X				
node.mme-name	X							
node.pgw-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

2.2.1.1.3 TCP

summary.ps_tcp-device-900-timeline

It contains information related to tcp protocol such as rtt or bearer packets for the devices (brand, model and type). The information is aggregated by device brand, device model, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.totalpackets				X				
bearer.totalpackets-samples				X				
device.brand	X							
device.model	X							
device.type	X							
imei.tac		X						
net.latency				X				
net.latency-samples				X				
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
tcp.retransmissions				X				
tcp.retransmissions-samples				X				
tcp.rtt-client				X				
tcp.rtt-client-samples				X				

tcp.rtt-server				X				
tcp.rtt-server-samples				X				
timestamp					X			

summary.ps_tcp-location-900-timeline

It contains information related to tcp protocol, aggregated by ENODB, cell, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.cell	X							
bearer.rat-type-str	X							
bearer.totalpackets				X				
bearer.totalpackets-samples				X				
cluster	X							
demographics.city	X							
demographics.region	X							
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
location.ci						X		
location.lac						X		
location.mcc						X		
location.mnc						X		
location.sac						X		
net.latency				X				
net.latency-samples				X				
node.enodeb-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
site-id	X							
tcp.retransmissions				X				
tcp.retransmissions-samples				X				
tcp.rtt-client				X				
tcp.rtt-client-samples				X				
tcp.rtt-server				X				
tcp.rtt-server-samples				X				
timestamp					X			

summary.ps_tcp-network-900-timeline

It contains information related to tcp protocol, aggregated by PGW, MME, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.totalpackets				X				
bearer.totalpackets-samples				X				
net.latency				X				
net.latency-samples				X				
node.mme-name	X							
node.pgw-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
tcp.retransmissions				X				
tcp.retransmissions-samples				X				
tcp.rtt-client				X				
tcp.rtt-client-samples				X				
tcp.rtt-server				X				
tcp.rtt-server-samples				X				
timestamp					X			

2.2.1.1.4 Web-browsing

summary.ps_browsing-device-900-timeline

It contains information related to browsing navigation, aggregated by device brand, device model, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
device.brand	X							
device.model	X							
device.type	X							
http.num-rsp				X				
http.num-rsp-4xx				X				
http.num-rsp-4xx-samples				X				
http.num-rsp-5xx				X				
http.num-rsp-5xx-samples				X				

http.num-rsp-samples				X				
http.session-time-duration					X			
http.session-time-duration-samples				X				
http.session-time-samples				X				
http.session-time-samples-samples				X				
imei.tac		X						
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.peak-throughputdown				X				
net.peak-throughputdown-samples				X				
net.peak-throughputup				X				
net.peak-throughputup-samples				X				
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

summary.ps_browsing-location-900-timeline

It contains information related to browsing navigation, aggregated by ENODB, cell, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.cell	X							
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
cluster	X							
demographics.city	X							
demographics.region	X							
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
http.num-rsp				X				
http.num-rsp-4xx				X				
http.num-rsp-4xx-samples				X				
http.num-rsp-5xx				X				
http.num-rsp-5xx-samples				X				
http.num-rsp-samples				X				

http.session-time-duration					X		
http.session-time-duration-samples				X			
http.session-time-samples				X			
http.session-time-samples-samples				X			
location.ci						X	
location.lac						X	
location.mcc						X	
location.mnc						X	
location.sac						X	
net.activity-duration-down							X
net.activity-duration-down-samples				X			
net.activity-duration-up							X
net.activity-duration-up-samples				X			
net.peak-throughputdown				X			
net.peak-throughputdown-samples				X			
net.peak-throughputup				X			
net.peak-throughputup-samples				X			
node.enodeb-name	X						
operator-country	X						
operator-id		X					
operator-name	X						
operator-tadig-code	X						
site-id	X						
timestamp					X		

summary.ps_browsing-network-900-timeline

It contains information related to browsing navigation, aggregated by PGW, MME, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
http.num-rsp				X				
http.num-rsp-4xx				X				
http.num-rsp-4xx-samples				X				
http.num-rsp-5xx				X				
http.num-rsp-5xx-samples				X				
http.num-rsp-samples				X				
http.session-time-duration					X			
http.session-time-duration-samples				X				
http.session-time-samples				X				

http.session-time-samples-samples				X				
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.peak-throughputdown				X				
net.peak-throughputdown-samples				X				
net.peak-throughputup				X				
net.peak-throughputup-samples				X				
node.mme-name	X							
node.pgw-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

summary.ps_browsing-url-900-timeline

It contains information related to browsing navigation, aggregated by URL, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.summary-total-activity				X				
bearer.summary-total-activity-samples				X				
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
http.activity-groups				X				
http.activity-groups-samples				X				
http.host				X				
http.num-rsp				X				
http.num-rsp-samples				X				
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

2.2.1.1.5 Streaming

summary.ps_streaming-device-900-timeline

It contains information related to video streaming such as rebuffering or reproduction time for the devices (brand, model and type). The information is aggregated by device brand, device model, video-codec and max-resolution.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
device.brand	X							
device.model	X							
device.type	X							
imei.tac		X						
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
stream.duration-samples				X				
stream.effective-reproduction-time								X
stream.effective-reproduction-time-samples				X				
stream.max-resolution				X				
stream.rebuffering-groups				X				
stream.rebuffering-groups-samples				X				
stream.reproduction-quality				X				
stream.reproduction-quality-samples				X				
stream.service-access-time								X
stream.service-access-time-samples				X				
stream.video-codec				X				
tcp.rebuffering-time								X
tcp.rebuffering-time-samples				X				
tcp.service-start-time					X			
tcp.service-start-time-samples				X				
timestamp					X			

summary.ps_streaming-network-900-timeline

It contains information related to video streaming such as rebuffering or reproduction time, aggregated by PGW, MME, video-codec and max-resolution.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
node.mme-name	X							

node.pgw-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
stream.duration-samples				X				
stream.effective-reproduction-time								X
stream.effective-reproduction-time-samples				X				
stream.max-resolution				X				
stream.rebuffering-groups				X				
stream.rebuffering-groups-samples				X				
stream.reproduction-quality				X				
stream.reproduction-quality-samples				X				
stream.service-access-time								X
stream.service-access-time-samples				X				
stream.video-codec				X				
tcp.rebuffering-time								X
tcp.rebuffering-time-samples				X				
tcp.service-start-time					X			
tcp.service-start-time-samples				X				
timestamp					X			

2.2.1.1.6 File-sharing

summary.ps_file_sharing-device-900-timeline

It contains information related to file sharing, aggregated by device brand, device model, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.summary-total-activity-down				X				
bearer.summary-total-activity-down-samples				X				
bearer.summary-total-activity-up				X				
bearer.summary-total-activity-up-samples				X				
bearer.total-volume-up				X				
bearer.total-volume-up-samples				X				
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
device.brand	X							
device.model	X							
device.type	X							
download-flow-samples				X				

flow-4xx-samples				X				
flow-5xx-samples				X				
flow-drop-samples				X				
flow-rst-samples				X				
flow-samples				X				
http.setup-time					X			
http.setup-time-samples				X				
imei.tac		X						
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.duration								X
net.duration-samples				X				
net.latency				X				
net.latency-samples				X				
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			
upload-flow-samples				X				

summary.ps_file_sharing-location-900-timeline

It contains information related to file sharing, aggregated by ENODB, cell, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.cell	X							
bearer.rat-type-str	X							
bearer.summary-total-activity-down				X				
bearer.summary-total-activity-down-samples				X				
bearer.summary-total-activity-up				X				
bearer.summary-total-activity-up-samples				X				
bearer.total-volume-up				X				
bearer.total-volume-up-samples				X				
bearer.volumedown				X				
bearer.volumedown-samples				X				
bearer.volumeup				X				
bearer.volumeup-samples				X				
cluster	X							
demographics.city	X							
demographics.region	X							
download-flow-samples				X				

flow-4xx-samples				X				
flow-5xx-samples				X				
flow-drop-samples				X				
flow-rst-samples				X				
flow-samples				X				
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
http.setup-time					X			
http.setup-time-samples				X				
location.ci						X		
location.lac						X		
location.mcc						X		
location.mnc						X		
location.sac						X		
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.duration								X
net.duration-samples				X				
net.latency				X				
net.latency-samples				X				
node.enodeb-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
site-id	X							
timestamp					X			
upload-flow-samples				X				

summary.ps_file_sharing-network-900-timeline

It contains information related to file sharing aggregated by PGW, MME, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
bearer.summary-total-activity-down				X				
bearer.summary-total-activity-down-samples				X				
bearer.summary-total-activity-up				X				
bearer.summary-total-activity-up-samples				X				
bearer.total-volume-up				X				
bearer.total-volume-up-samples				X				
bearer.volumedown				X				
bearer.volumedown-samples				X				

bearer.volumeup				X				
bearer.volumeup-samples				X				
download-flow-samples				X				
flow-4xx-samples				X				
flow-5xx-samples				X				
flow-drop-samples				X				
flow-rst-samples				X				
flow-samples				X				
http.setup-time					X			
http.setup-time-samples				X				
net.activity-duration-down								X
net.activity-duration-down-samples				X				
net.activity-duration-up								X
net.activity-duration-up-samples				X				
net.duration								X
net.duration-samples				X				
net.latency				X				
net.latency-samples				X				
node.mme-name	X							
node.pgw-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			
upload-flow-samples				X				

2.2.1.2 Control-Plane

These data sources contain control-plane information that create the signaling to manage the traffic.

2.2.1.2.1 Signaling (Gn/S11)

summary.ps_signaling-apn-900-timeline

It contains information related to signaling records such as pdp creation time or failure ratio for each APN. The information is aggregated by APN and rat-type.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
bearer.apn		X						
bearer.rat-type-str	X							
create-bearer-failure-samples				X				
create-bearer-response-samples				X				

create-pdp-failure-samples				X				
create-pdp-response-samples				X				
create-pdp-time								X
create-pdp-time-samples				X				
create-session-failure-samples				X				
create-session-response-samples				X				
create-session-time								X
create-session-time-samples				X				
delete-pdp-network-request-samples				X				
delete-pdp-request-samples				X				
modify-bearer-failure-samples				X				
modify-bearer-response-samples				X				

summary.ps_signaling-device-900-timeline

It contains information related to signalling records such as pdp creation time or failure ratio for devices (brand, model and type). The information is aggregated by device brand, device model, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
create-bearer-failure-samples				X				
create-bearer-response-samples				X				
create-pdp-failure-samples				X				
create-pdp-response-samples				X				
create-pdp-time								X
create-pdp-time-samples				X				
create-session-failure-samples				X				
create-session-response-samples				X				
create-session-time								X
create-session-time-samples				X				
delete-pdp-network-request-samples				X				
delete-pdp-request-samples				X				
device.brand	X							
device.model	X							
device.type	X							
imei.tac		X						
modify-bearer-failure-samples				X				
modify-bearer-response-samples				X				
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

summary.ps_signaling-location-900-timeline

It contains information related to signalling records such as pdp creation time or failure ratio for each location at cell level. The information is aggregated by ENODB, cell, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.cell	X							
bearer.rat-type-str	X							
cluster	X							
create-bearer-failure-samples				X				
create-bearer-response-samples				X				
create-pdp-failure-samples				X				
create-pdp-response-samples				X				
create-pdp-time								X
create-pdp-time-samples				X				
create-session-failure-samples				X				
create-session-response-samples				X				
create-session-time								X
create-session-time-samples				X				
delete-pdp-network-request-samples				X				
delete-pdp-request-samples				X				
demographics.city	X							
demographics.region	X							
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
location.ci						X		
location.lac						X		
location.mcc						X		
location.mnc						X		
location.sac						X		
modify-bearer-failure-samples				X				
modify-bearer-response-samples				X				
node.enodeb-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
site-id	X							
timestamp					X			

summary.ps_signaling-network-900-timeline

It contains information related to signaling records such as pdp creation time or failure ratio, aggregated by PGW, MME, rat-type and operator.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
be-mvno	X							
bearer.rat-type-str	X							
create-bearer-failure-samples				X				
create-bearer-response-samples				X				
create-pdp-failure-samples				X				
create-pdp-response-samples				X				
create-pdp-time								X
create-pdp-time-samples				X				
create-session-failure-samples				X				
create-session-response-samples				X				
create-session-time								X
create-session-time-samples				X				
delete-pdp-network-request-samples				X				
delete-pdp-request-samples				X				
modify-bearer-failure-samples				X				
modify-bearer-response-samples				X				
node.mme-name	X							
node.pgw-name	X							
operator-country	X							
operator-id		X						
operator-name	X							
operator-tadig-code	X							
timestamp					X			

2.3 Data sources for raw XDRs

In these data sources the information is inserted such as it comes from the probes: some decoration rules and conditional filters could be applied if necessary but the information is available as you could see it at probe level. Therefore the drilldown to subscriber data is allowed in this kind of data source.

The information is available in the data stores in about 15 minutes from the data is generated. And this data is stored in the BBDD for around 15 days, depending on the client needs and the HW resources availability.

Current data source list:

- datastore.ps_browsing_records

- datastore.ps_dns_records
- datastore.ps_file_sharing_records
- datastore.ps_session_records
- datastore.ps_signaling_records
- datastore.ps_streaming_records
- datastore.ps_tcp_records
- datastore.crm_records

The following sections depict the available tables of this data source and list all the fields for each table.

Check the appendix of this document for a further explanation of each field.

2.3.1 PS/Zhilabs

The following data sources store data that has been extracted from Zhilabs App level probes.

2.3.1.1 User-Plane

These data sources contain user-plane information related to data and app traffic.

2.3.1.1.1 Session

It refers to events that are produced in the period of time in which the user is connected maintaining active the session

datastore.ps_session_records

It contains information related to sessions such as session volume or throughput per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field
app-consent-level	X							X	
be-mvno	X								
bearer.apn		X							
bearer.cell	X								
bearer.charging-id		X							
bearer.imeisv		X							
bearer.imsi			X						
bearer.msisdn			X						
bearer.packetsdown				X					
bearer.packetsup				X					
bearer.rat-type-str	X								
bearer.start-time					X				
bearer.stop-time					X				
bearer.summary-average-throughput				X					

bearer.summary-average-throughput-down			X			
bearer.summary-average-throughput-up			X			
bearer.summary-throughput			X			
bearer.summary-throughput-down			X			
bearer.summary-throughput-up			X			
bearer.summary-total-activity			X			
bearer.summary-total-activity-down			X			
bearer.summary-total-activity-up			X			
bearer.total-volume-down			X			
bearer.total-volume-up			X			
bearer.totalpackets			X			
bearer.totalvolume			X			
bearer.user-ip		X				
bearer.user-ip6		X				
bearer.volumedown			X			
bearer.volumeup			X			
cell-consent-level	X					X
cluster	X					
crm-index					X	
demographics.city	X					
demographics.region	X					
device.brand	X					
device.model	X					
device.type	X					
event.interim-id		X				
event.start-time				X		
event.stop-time				X		
event.text			X			
gtp.ecgi-cellid					X	
gtp.ecgi-enbid					X	
gtp.ecgi-mcc					X	
gtp.ecgi-mnc					X	
gtp.ggsn-c		X				
gtp.ggsn-u		X				
gtp.qos-max-dl-rate-bps			X			
gtp.qos-max-ul-rate-bps			X			
gtp.rai-lac					X	
gtp.rai-mcc					X	
gtp.rai-mnc					X	
gtp.rai-rac					X	
gtp.rat-type			X			
gtp.session-id			X			
gtp.sgsn-c		X				
gtp.sgsn-u		X				
imei.tac		X				
location.ci					X	
location.lac					X	
location.mcc					X	

location.mnc					X				
location.sac					X				
lte.enodeb					X				
lte.mme					X				
lte.pgw-c		X							
lte.pgw-u		X							
lte.sgw-c		X							
lte.sgw-u		X							
net.activity-duration									X
net.activity-duration-down									X
net.activity-duration-up									X
net.duration									X
net.effective-throughput				X					
net.effective-throughputdown				X					
net.effective-throughputup				X					
net.peak-throughputdown				X					
net.peak-throughputup				X					
net.throughputdown				X					
net.throughputup				X					
net.uplink				X					
node.enodeb-name	X								
node.mme-name	X								
node.pgw-name	X								
operator-country	X								
operator-id		X							
operator-name	X								
operator-tadig-code	X								
primary_offering							X		
site-id	X								
sub-type							X		
subscriber.duration-usage									X
timestamp						X			

2.3.1.1.2 DNS

A DNS record is generated when an application used by the subscriber makes a enquiry to the server in order to resolve a domain name

datastore.ps_dns_records

It contains information related to DNS operations such as DNS codes or latency per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field

app-consent-level	X						X	
be-mvno	X							
bearer.apn		X						
bearer.cell	X							
bearer.imeisv		X						
bearer.imsi			X					
bearer.msisdn			X					
bearer.rat-type-str	X							
bearer.start-time					X			
bearer.stop-time					X			
cell-consent-level	X						X	
cluster	X							
demographics.city	X							
demographics.region	X							
device.brand	X							
device.model	X							
device.type	X							
dns.code		X						
event.interim-id		X						
event.start-time					X			
event.stop-time					X			
event.text				X				
gtp.ecgi-cellid						X		
gtp.ecgi-enbid						X		
gtp.ecgi-mcc						X		
gtp.ecgi-mnc						X		
gtp.ggsn-c		X						
gtp.ggsn-u		X						
gtp.rat-type				X				
gtp.sgsn-c		X						
gtp.sgsn-u		X						
imei.tac		X						
location.ci						X		
location.lac						X		
location.mcc						X		
location.mnc						X		
location.sac						X		
lte.enodeb						X		
lte.mme						X		
lte.pgw-c		X						
lte.pgw-u		X						
lte.sgw-c		X						
lte.sgw-u		X						
net.cell-ip				X				
net.cell-port				X				
net.duration								X
net.latency				X				
net.srv-ip				X				

net.srv-port			X				
node.enodeb-name	X						
node.mme-name	X						
node.pgw-name	X						
operator-country	X						
operator-id		X					
operator-name	X						
operator-tadig-code	X						
site-id	X						
timestamp				X			
segment						X	
subscriber_status						X	
gtp.packet_1			X				
location.mcc					X		
location.mnc					X		
location.sac					X		
lte.enodeb					X		
lte.mme					X		
lte.pgw-c		X					
lte.pgw-u		X					
lte.sgw-c		X					
lte.sgw-u		X					
net.activity-duration							X
net.activity-duration-down							X
net.activity-duration-up							X
net.duration							X
net.effective-throughput			X				
net.effective-throughputdown			X				
net.effective-throughputup			X				
net.peak-throughputdown			X				
net.peak-throughputup			X				
net.throughputdown			X				
net.throughputup			X				
net.uplink			X				
node.enodeb-name	X						
node.mme-name	X						
node.pgw-name	X						
operator-country	X						
operator-id		X					
operator-name	X						
operator-tadig-code	X						
primary_offering						X	
site-id	X						
sub-type						X	
subscriber.duration-usage							X
timestamp				X			

2.3.1.1.3 TCP

A TCP record is generated when an application used by the subscriber makes a tcp transaction

datastore.ps_tcp_records

It contains information related to tcp protocol such as rtt or retransmissions per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field
app-consent-level	X							X	
app.category	X								
app.name				X					
be-mvno	X								
bearer.apn		X							
bearer.cell	X								
bearer.imsi			X						
bearer.msisdn			X						
bearer.rat-type-str	X								
bearer.start-time					X				
bearer.stop-time					X				
bearer.totalpackets				X					
bearer.totalvolume				X					
bearer.volumedown				X					
bearer.volumeup				X					
cell-consent-level	X							X	
cluster	X								
demographics.city	X								
demographics.region	X								
device.brand	X								
device.model	X								
device.type	X								
event.interim-id		X							
event.start-time					X				
event.stop-time					X				
event.text				X					
gtp.ecgi-cellid						X			
gtp.ecgi-enbid						X			
gtp.ecgi-mcc						X			
gtp.ecgi-mnc						X			
gtp.ggsn-u		X							
gtp.qos-max-dl-rate-bps				X					
gtp.qos-max-ul-rate-bps				X					
gtp.sgsn-c		X							
gtp.sgsn-u		X							
http.code				X					
http.content-type				X					

http.host			X			
http.session-time				X		
http.setup-time				X		
http.user-agent			X			
imei.tac		X				
location.ci					X	
location.lac					X	
location.mcc					X	
location.mnc					X	
location.sac					X	
lte.mme					X	
lte.pgw-u		X				
lte.sgw-u		X				
net.activity-duration						X
net.activity-duration-down						X
net.activity-duration-up						X
net.cell-ip			X			
net.duration						X
net.latency			X			
net.peak-throughputdown			X			
net.peak-throughputup			X			
net.srv-ip			X			
net.tag			X			
net.termination-code			X			
net.throughputdown			X			
net.throughputup			X			
node.enodeb-name	X					
node.mme-name	X					
node.pgw-name	X					
operator-country	X					
operator-id		X				
operator-name	X					
operator-tadig-code	X					
site-id	X					
tcp.retransmissions			X			
tcp.rtt-client			X			
tcp.rtt-server			X			
timestamp				X		

2.3.1.1.4 Web-browsing

A web-browsing record is generated when the subscriber browses the internet using a browsing application

datastore.ps_browsing_records

It contains information related to browsing navigation such as http volume or session time per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field
app-consent-level	X							X	
be-mvno	X								
bearer.apn		X							
bearer.cell	X								
bearer.charging-id		X							
bearer.imeisv		X							
bearer.imsi			X						
bearer.msisdn			X						
bearer.packetsdown				X					
bearer.packetsup				X					
bearer.rat-type-str	X								
bearer.start-time					X				
bearer.stop-time					X				
bearer.summary-average-throughput				X					
bearer.summary-average-throughput-down				X					
bearer.summary-average-throughput-up				X					
bearer.summary-throughput				X					
bearer.summary-throughput-down				X					
bearer.summary-throughput-up				X					
bearer.summary-total-activity				X					
bearer.summary-total-activity-down				X					
bearer.summary-total-activity-up				X					
bearer.total-volume-down				X					
bearer.total-volume-up				X					
bearer.totalpackets				X					
bearer.totalvolume				X					
bearer.user-ip		X							
bearer.user-ip6		X							
bearer.volumedown				X					
bearer.volumeup				X					
cell-consent-level	X							X	
cluster	X								
demographics.city	X								
demographics.region	X								
device.brand	X								
device.model	X								
device.type	X								
event.interim-id		X							
event.start-time					X				
event.stop-time					X				
event.text				X					
gtp.ecgi-cellid						X			
gtp.ecgi-enbid						X			
gtp.ecgi-mcc						X			
gtp.ecgi-mnc						X			

gtp.ggsn-c		X			
gtp.ggsn-u		X			
gtp.qos-max-dl-rate-bps			X		
gtp.qos-max-ul-rate-bps			X		
gtp.rai-lac				X	
gtp.rai-mcc				X	
gtp.rai-mnc				X	
gtp.rai-rac				X	
gtp.rat-type			X		
gtp.session-id			X		
gtp.sgsn-c		X			
gtp.sgsn-u		X			
http.activity-groups			X		
http.host			X		
http.num-rsp			X		
http.num-rsp-4xx			X		
http.num-rsp-5xx			X		
http.session-time-activity				X	
http.session-time-duration				X	
http.session-time-samples			X		
imei.tac		X			
location.ci				X	
location.lac				X	
location.mcc				X	
location.mnc				X	
location.sac				X	
lte.enodeb				X	
lte.mme				X	
lte.pgw-c		X			
lte.pgw-u		X			
lte.sgw-c		X			
lte.sgw-u		X			
net.activity-duration					X
net.activity-duration-down					X
net.activity-duration-up					X
net.duration					X
net.effective-throughput			X		
net.effective-throughputdown			X		
net.effective-throughputup			X		
net.peak-throughputdown			X		
net.peak-throughputup			X		
net.throughputdown			X		
net.throughputup			X		
net.uplink			X		
node.enodeb-name	X				
node.mme-name	X				
node.pgw-name	X				
operator-country	X				

operator-id		X							
operator-name	X								
operator-tadig-code	X								
site-id	X								
subscriber.duration-usage									X
timestamp					X				
bearer.imeisv		X							

2.3.1.1.5 Streaming

A streaming record is generated when the subscriber consumes streaming services in the device

datastore.ps_streaming_records

It contains information related to video streaming such as rebuffering or video resolution per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field
app-consent-level	X							X	
app.name				X					
be-mvno	X								
bearer.apn		X							
bearer.cell	X								
bearer.imsi			X						
bearer.msisdn			X						
bearer.rat-type-str	X								
bearer.start-time					X				
bearer.stop-time					X				
bearer.totalpackets				X					
bearer.totalvolume				X					
bearer.volumedown				X					
bearer.volumeup				X					
cell-consent-level	X								
cluster	X							X	
demographics.city	X								
demographics.region	X								
device.brand	X								
device.model	X								
device.type	X								
event.interim-id		X							
event.start-time					X				
event.stop-time					X				
event.text				X					
gtp.ecgi-cellid						X			
gtp.ecgi-enbid						X			

gtp.ecgi-mcc					X	
gtp.ecgi-mnc					X	
gtp.ggsn-u		X				
gtp.qos-max-dl-rate-bps			X			
gtp.qos-max-ul-rate-bps			X			
gtp.sgsn-c		X				
gtp.sgsn-u		X				
http.code			X			
http.content-type			X			
http.host			X			
http.session-time				X		
http.setup-time				X		
http.user-agent			X			
imei.tac		X				
location.ci					X	
location.lac					X	
location.mcc					X	
location.mnc					X	
location.sac					X	
lte.mme					X	
lte.pgw-u		X				
lte.sgw-u		X				
net.activity-duration						X
net.activity-duration-down						X
net.activity-duration-up						X
net.cell-ip			X			
net.duration						X
net.latency			X			
net.peak-throughputdown			X			
net.peak-throughputup			X			
net.srv-ip			X			
net.tag			X			
net.termination-code			X			
net.throughputdown			X			
net.throughputup			X			
node.enodeb-name	X					
node.mme-name	X					
node.pgw-name	X					
operator-country	X					
operator-id		X				
operator-name	X					
operator-tadig-code	X					
site-id	X					
stream.audio-codec			X			
stream.duration						X
stream.effective-reproduction-time						X
stream.max-resolution			X			
stream.rebuffering-groups			X			

stream.reproduction-quality				X				
stream.service-access-time					X			
stream.video-codec				X				
stream.video-resolutions				X				
stream.video-resolutions-time					X			
tcp.rebuffering-time								X
tcp.retransmissions				X				
tcp.service-start-time					X			
timestamp					X			

2.3.1.1.6 File-sharing

A file-sharing record is generated when the subscriber uses a sharing service in which files exchanges are made

datastore.ps_file_sharing_records

It contains information related to file sharing such as volume or http service setup time per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field
app-consent-level	X							X	
app.category	X								
app.name				X					
be-mvno	X								
bearer.apn		X							
bearer.cell	X								
bearer.imsi			X						
bearer.msisdn			X						
bearer.rat-type-str	X								
bearer.start-time					X				
bearer.stop-time					X				
bearer.summary-total-activity-down				X					
bearer.summary-total-activity-up				X					
bearer.total-volume-down				X					
bearer.total-volume-up				X					
bearer.totalpackets				X					
bearer.totalvolume				X					
bearer.volumedown				X					
bearer.volumeup				X					
cell-consent-level	X							X	
cluster	X								
demographics.city	X								
demographics.region	X								
device.brand	X								

device.model	X					
device.type	X					
event.interim-id		X				
event.start-time				X		
event.stop-time				X		
event.text			X			
gtp.ecgi-cellid					X	
gtp.ecgi-enbid					X	
gtp.ecgi-mcc					X	
gtp.ecgi-mnc					X	
gtp.ggsn-u		X				
gtp.qos-max-dl-rate-bps			X			
gtp.qos-max-ul-rate-bps			X			
gtp.sgsn-c		X				
gtp.sgsn-u		X				
http.code			X			
http.content-type			X			
http.host			X			
http.session-time				X		
http.setup-time				X		
http.user-agent			X			
imei.tac		X				
location.ci					X	
location.lac					X	
location.mcc					X	
location.mnc					X	
location.sac					X	
lte.mme					X	
lte.pgw-u		X				
lte.sgw-u		X				
net.activity-duration						X
net.activity-duration-down						X
net.activity-duration-up						X
net.cell-ip			X			
net.duration						X
net.latency			X			
net.peak-throughputdown			X			
net.peak-throughputup			X			
net.srv-ip			X			
net.tag			X			
net.termination-code			X			
net.throughputdown			X			
net.throughputup			X			
node.enodeb-name	X					
node.mme-name	X					
node.pgw-name	X					
operator-country	X					
operator-id		X				

operator-name	X							
operator-tadig-code	X							
site-id	X							
tcp.retransmissions				X				
tcp.rtt-client				X				
tcp.rtt-server				X				
timestamp					X			

2.3.1.2 Control-Plane

These data sources contain control-plane information that create the signaling to manage the traffic.

2.3.1.2.1 Signaling (Gn/S11)

Signal events that are produced to control the communication during the period in which the subscriber is connected

datastore.ps_signaling_records

It contains information related to signalling records such as the signalling events per subscriber. There is no aggregation and the information is stored in raw.

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Privacy	Duration Field
app-consent-level	X							X	
be-mvno	X								
bearer.apn		X							
bearer.cell	X								
bearer.charging-id		X							
bearer.imeisv		X							
bearer.imsi			X						
bearer.msisdn			X						
bearer.rat-type-str	X								
bearer.user-ip		X							
bearer.user-ip6		X							
cell-consent-level	X							X	
cluster	X								
demographics.city	X								
demographics.region	X								
device.brand	X								
device.model	X								
device.type	X								
event.interim-id		X							
event.start-time					X				
event.stop-time					X				
event.text				X					
gtp.cause				X					

gtp.dst			X			
gtp.duration						X
gtp.ecgi-cellid				X		
gtp.ecgi-enbid				X		
gtp.ecgi-mcc				X		
gtp.ecgi-mnc				X		
gtp.ggsn-c		X				
gtp.ggsn-u		X				
gtp.latency			X			
gtp.packet			X			
gtp.qos-max-dl-rate-bps			X			
gtp.qos-max-ul-rate-bps			X			
gtp.rai-lac				X		
gtp.rai-mcc				X		
gtp.rai-mnc				X		
gtp.rai-rac				X		
gtp.rat-type			X			
gtp.session-id			X			
gtp.sgsn-c		X				
gtp.sgsn-u		X				
gtp.src			X			
gtp.version			X			
imei.tac		X				
location.ci				X		
location.lac				X		
location.mcc				X		
location.mnc				X		
location.sac				X		
lte.enodeb				X		
lte.mme				X		
lte.pgw-c		X				
lte.pgw-u		X				
lte.sgw-c		X				
lte.sgw-u		X				
net.tag			X			
net.uplink			X			
node.dst-name	X					
node.enodeb-name	X					
node.mme-name	X					
node.pgw-name	X					
node.src-name	X					
operator-country	X					
operator-id		X				
operator-name	X					
operator-tadig-code	X					
site-id	X					
timestamp					X	

2.3.1.3 CRM

It contains information related to customer relationship management. There is no aggregation and the information is stored in raw

FIELDS	Decoration	From Network	Subscriber Identifier	Indicator	Timestamp	Location Field	CRM Field	Duration Field
activation_date							X	
imsi			X					
last-month_revenue							X	
msisdn			X					
nationality							X	
primary_offering							X	
race							X	
segment							X	
sim_type							X	
sub-type							X	
subscriber_status							X	
supplementary_offering							X	
timestamp					X			

Chapter 3
Functions

In this section, we will learn to write queries that can be executed in the Query Console. The section starts with the simplest queries and finishes with the most complex ones. You need to practise a little bit, but the bases are quite simple and advancing in complexity becomes easy once you understand the rules.

In addition, you are not alone, as the CEA provides you an on-line help with auto completion function related to the section of the query you are writing.

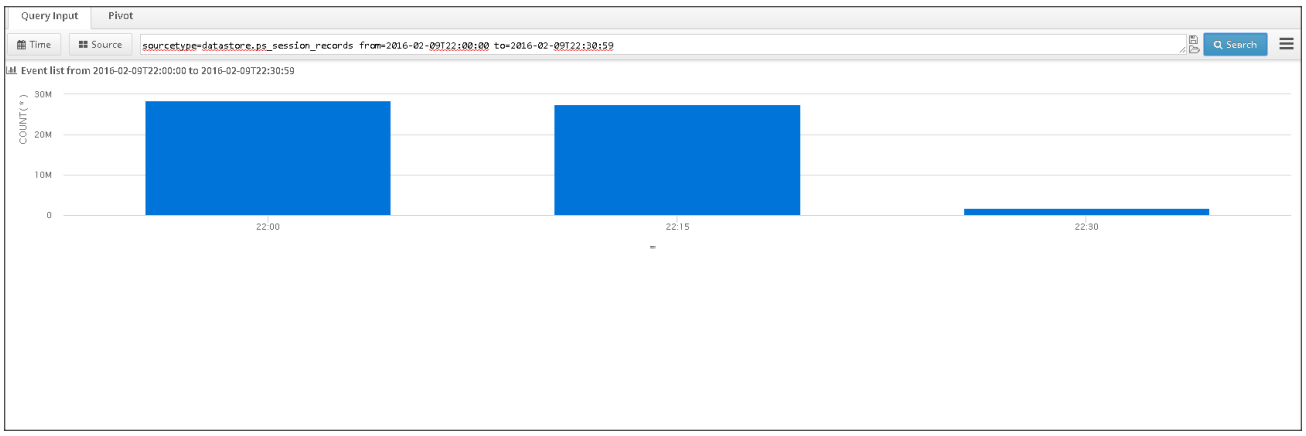
Let's start with the base. In a summarized way a query is always composed of these elements:

- (M) Data Range
- (M) Data Source (Source Type)
- (O) Aggregation Period
- (O) Summarization Function
- (O) Filters
- (O) Functions

Where (M) are mandatory and (O) are optional. This means that the simplest query is that one composed by a Data Range and a Data Source.

3.1 Data Range and Data Source

The simplest CEA Query requires at least a Data Range and a Data Source. A query containing these two elements provides as a result the temporal evolution of the count of all the events contained in the data source for the selected time range which is an Event Time Line chart. It is very usual to start the high level inspection with this kind of queries to detect for *example* abnormal values.

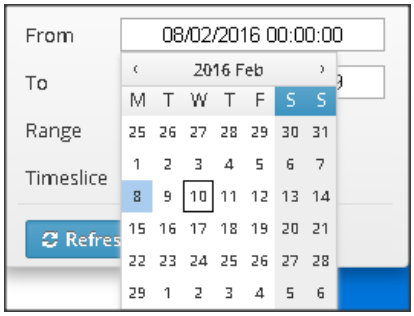


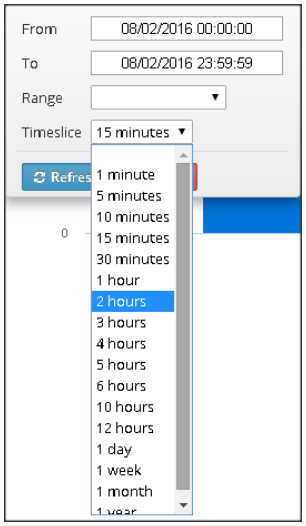
3.1.1 Configuring Data Range and Temporal Aggregation

Data Range is selected by means of the Data Range Selector or, alternatively, it can be written directly in the Query Command Line in the format *YYYY-mm-ddTHH:MM:SS*.

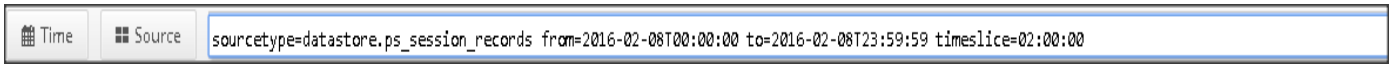
The *Data Range* can be a fixed from – to date in the format *YYYY-mm-ddTHH:MM:SS* that you can directly pick and edit in the Data Range Selector, but it can also be a relative range as for *example* last day or last week.

Additionally, the Data Range Selector allow the selection of a *Timeslice* which is an optional parameter indicating the temporal aggregation of data for the query calculation.

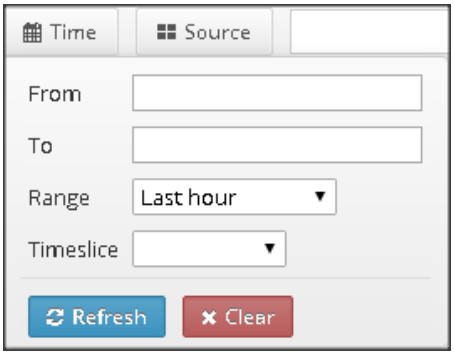




By clicking the Refresh Button, the selected Data Range and timeslice will be automatically written in the Command Line.



You also can choose, last hour, last day… by selecting the option in "Range"

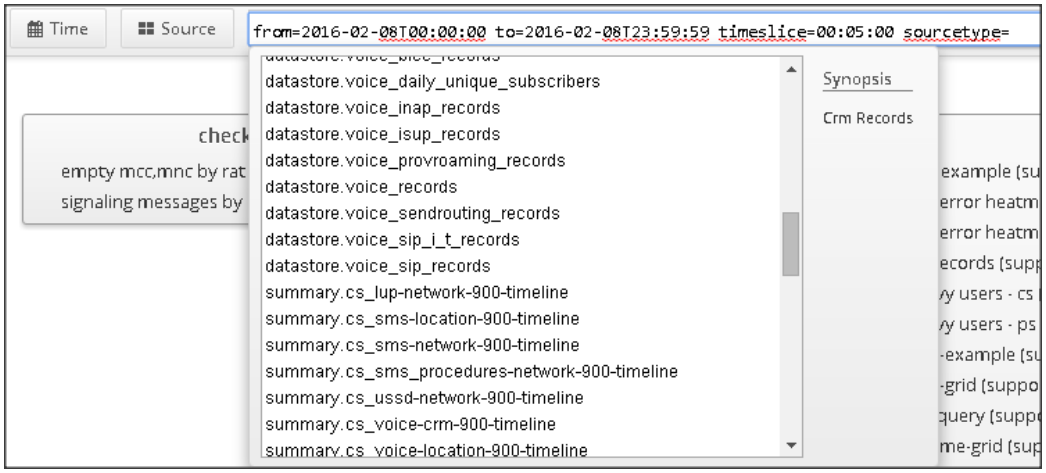


3.2 Selecting a data source

The data source is each of the different kind of data the CEA is feed of. Data source selection is done by typing:

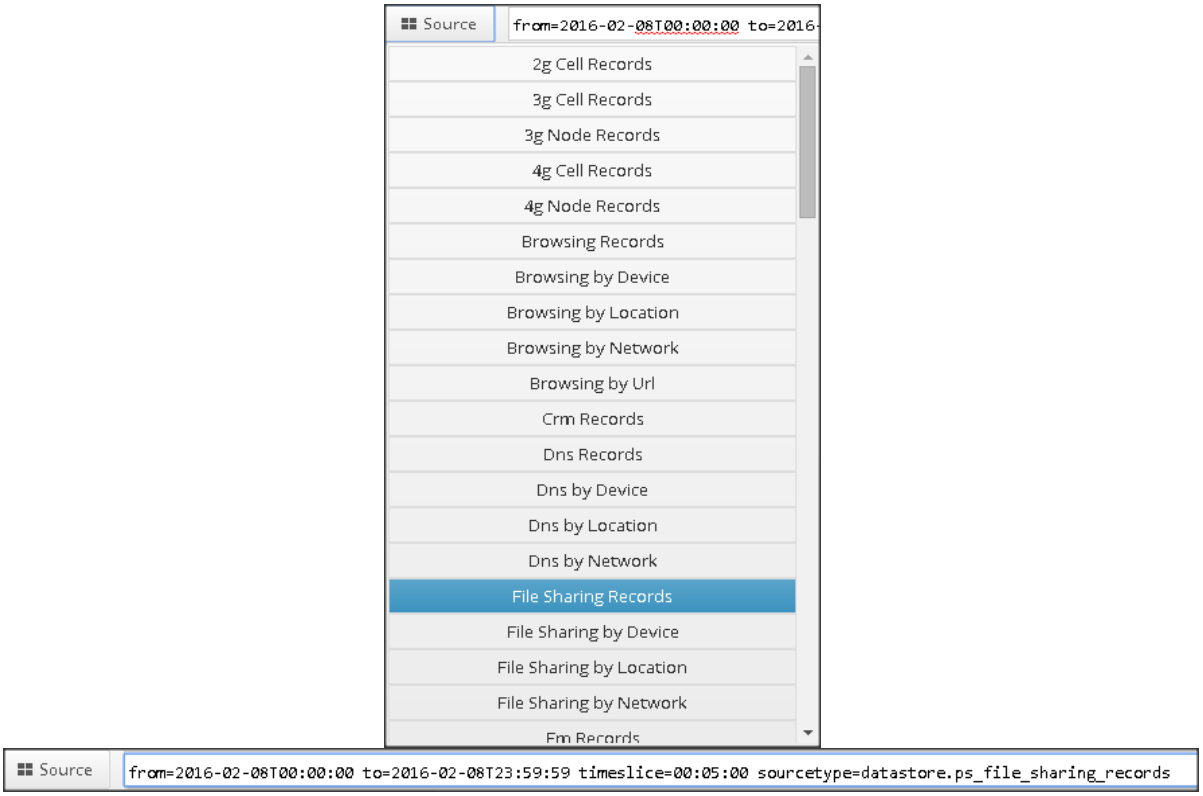
sourcetype=<datastore_name>

CEA includes an on-line help so that when you write the equal symbol the list of available data sources will be displayed making the selection easier by just clicking on the one you are interested in.



Please, note that the list of available data sources depends on the acquired licence and the integrated data sources.

You also can click on “Source” button and the list with all data sources available will be displayed



3.3 Sampling

The sampling function is an optional feature that applies a sampling rate in the computation of the query providing a faster response for the user to get a quick view on the result.

The syntax of sampling is as simple as including in *sampling=<sampling rate>*

A sampling rate equal to 10 means that the system will take 1 of every 10 samples to compute the query.

Example:

From XXX to XXXX data source=sourcetype… sampling=10

Sampling for Full Search queries in CEA is designed to reduce query response time and system load for non-subscriber specific queries executed over large records tables. Technically, setting sampling=n in the header section of a CEA Full Search query, it slices the persisted data store such as to query only every n=-th slice of data. Results are adjusted by a factor of n for count and sum aggregations only. Sampling accuracy relies on number of slices, and sample set size and statistical distribution. The smaller the result buckets, the less accurate the sampling result.

Note: Sampling is typically not efficient for small tables, including time aggregated summary, and small result buckets, typical for unique counts. Advice is to validate sampling efficiency per query and data set.

3.4 Filtering by a subscriber or subscribers

It is possible to filter the registers to a specific subscriber or a list of them. To do so it is necessary to add the MSISDN(s) of the subscriber(s) after the sourcetype and the query will be restrained to that particular subscriber(s) separated by space.

Syntax:

From=<from_date> to=<to_date> sourcetype=<sourcetype_name> <SUBSCRIBER_MSISDN1> <SUBSCIBER_MSISDN2> … <SUBSCRIBER_MSISDNN>

Example:

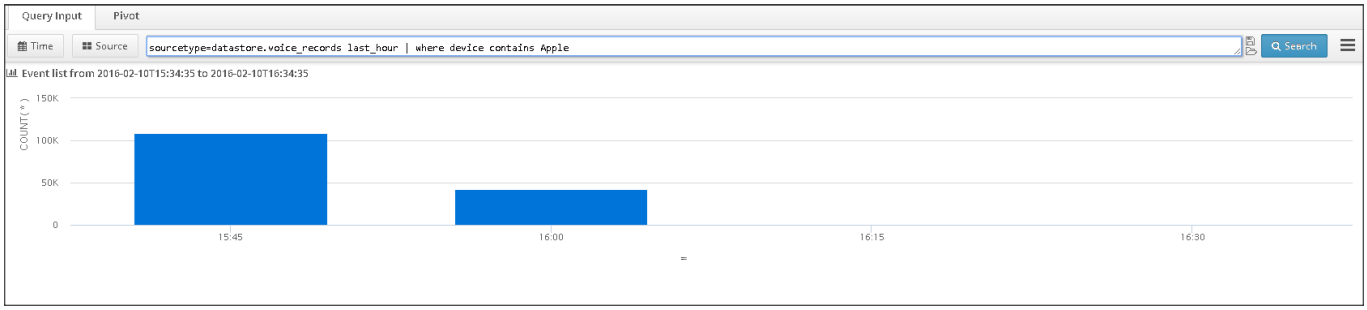
From=<from_date> to=<to_date> data source=datastore.volte_sip 999111999

Note: In some raw data sources the key for querying it is the IMSI instead of the MSISDN. It depends on the database implementation so the previous paragraph would apply to the corresponding key in each case.

3.5 Filters

Filters introduce a condition that the data displayed on the result of the query will always obey. They are optional and it is possible to add from 0 to N filters on a specific query. The addition of more than one filter behaves as an AND operation on the data. This means that if, for *example*, there are two conditions filtered, it is necessary that both of them are matched on a register in order to consider it on the query result.

Filters are separated from the first part of the query by a "|"



***Syntax*:**

From=<from_date> to=<to_date> sourcetype=<sourcetype_name>|where <field> <operator> <value>

<operator> It can be one of the following:

Table 1. Operators

Operator	Meaning
<	Less than
>	Greater than
<=	Less than or equal
>=	Greater than or equal
=	Equal
!=	Not equal
starts-with	Starts with for strings
!starts with	Does not starts with for strings
ends-with	Ends with for strings
!ends-with	Does not end with for strings
contains	Contains for strings
!contains	Does not contain for string

<*value*> It can be a number or a string. It depends on the field to filter

Example 1:

sourcetype=datastore.ps_file_sharing_records last_hour | where bearer.volumeup > 300

Example 2:

sourcetype=datastore.ps_streaming_records from=2016-01-22T00:00:00 to=2016-01-22T23:59:59 | where http.session-time > 2790| where device.type !contains "Mobile Handset"

Example 3:

from=2016-01-22T00:00:00 to=2016-01-22T23:59:59 sourcetype= datastore.ps_streaming_records | where device contains samsung

3.6 Functions

The use of functions provides an extremely flexibility and power to the CEA Query Language. This section describes them with different *examples* for a better understanding of them.

3.6.1 chart

The chart function refers to the graphical representation of the data. There is a wide range of available charts each of one better suit the representation of a query result.

Syntax

chart type field [fieldN] by field_By [field_ByN] [-height=num] [-width=num] [-legend=no] [-ylabel=no] [-layout=tab|horizontal|vertical|gridWxH] [-override-title=yes] [-table=no]

[] means they are optional parameters, otherwise they are obligatory.

So, the simplest chart query is:

chart type field by field_By

Parameters

type can be one of the following elements:

Table 2. <Type> parameter description

<Type>	Meaning
pie	Displays the diagram as a pie chart
pie3d	Displays the diagram as a 3D pie chart
stackedpie	Displays the diagram as a stacked donut chart
donut	Displays the diagram as a donut chart
donut3d	Displays the diagram as a 3D donut chart
column or col	Displays the diagram as a column chart
column3d or col3d	Displays the diagram as a 3D column chart
bar	Displays the diagram as a bar chart
bar3d	Displays the diagram as a 3D bar chart
stackedbar	Displays the diagram as a stacked bar chart
stackedcol	Displays the diagram as a stacked col chart
treemap	Displays the diagram as a treemap chart
bubble	Displays the diagram as a bubble chart
timeline	Displays the diagram as a timeline with the values in chronological order. Supersedes dygraph
dygraph	Displays the diagram as timeline with the values in chronological order
stackedtimeline	Displays the diagram as a timeline with the values in chronological order with stacked values
filledareatimeline	Displays the diagram as a timeline with the values in chronological order with filled area
stackedareatimeline	Displays the diagram as a timeline with the values in chronological order with stacked values and filled area
stackedtimebar, timebar	Displays the diagram as stacked bars in chronological order
groupedtimebar	Displays the diagram as grouped bars with the values in chronological order
map	Displays the diagram as a geographical representation of the location information in the records
table	Displays the diagram as an exportable table
summary-table	Displays the diagram as an exportable summarized table
speedometer	Displays the diagram as a speedometer. It needs to define max limit value("-max-limit="), and high("-medium-high=") and low("-medium-low=") values for the medium range. • For indicators which high values are ok and low values are not ok configure the speedometer using: max limit

	value > medium-high value > medium-low. In this case: – the range from 0 to medium-low is painted in red. – the range from medium-low to medium-high is painted in yellow. – the range from medium-high to max limit value is painted in green. • For indicators which high values are nok and low values are ok configure the speedometer using: max limit value > medium-low >medium-high value: – the range from 0 to medium-high is painted in green. – the range from medium-high to medium-low is painted in yellow – the range from medium-low to max limit value is painted in green. At the end of the section there are some *examples* with speedometer chart.
gauge	Displays the diagram as a simple gauge visialization
label	Displays the result in a text
filledareatimeline	Displays the result ina a filled area timeline
trend-label	Displays the result in trend chart summary
kiviat	Displays the result in a kiviat/radar diagram
polar	Displays the result in a polar column diagram
scatter2	Displays the result in a simple scatter plot

field [fieldN] can be one of the following elements:

- One of the fields contained in the query sourcetype.
- A function included in Table 3.

Table 3. <Function> Parameter Description

<function>	**Meaning**
avg(fieldA)	This function calculates the average value of the field specified in fieldA. Where fieldA is one of fields contained in the sourcetype.
median(fieldA)	This function calculates the median value of the field specified in fieldA
max(fieldA)	This function calculates the maximun value of the field specified in fieldA
min(fieldA)	This function calculates the minimun value of the field specified in fieldA
sum(fieldA)	This function sum the values of the field specified in fieldA

count(fieldA)	This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified
unique(fieldA)	This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified
avg_if(fieldA, expression)	This function calculates the average value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
median_if(fieldA, expression)	This function calculates the median value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
max_if(fieldA, expression).	This function calculates the maximun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
max_if(fieldA, expression).	This function calculates the maximun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
min_if(fieldA, expression).	This function calculates the minimun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
sum_if(fieldA, expression)	This function sum the values of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
count_if(fieldA, expression)	This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified when the expression defined is true. Available expressions are the same as the ones used in where function
unique_if(fieldA, expression)	This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified when the expression defined is true. Available expressions are the same as the ones used in where function

- *[fieldA operator fieldB] [as expressionName]* which is an expression or group of expressions in brackets, where:
  - *fieldA, fieldB* can be:
    - one of the fields included in sourcetype.
    - numerical values
    - a function as described in Table 3.
  - *operator* can be any of the following arithmetic operators included in Table 4.
  - *as expressionName* it is an optional parameter to specify an alias for the expression

Table 4. <Operator> Parameter Description

<operator>	Meaning
+	Addition
-	Subtraction
*	Multiplication
/	Division

Examples:

... | *chart pie [acct-output-octets + acct-input-octets] as volume by rat-type*

... | *chart pie [acct-output-octets + acct-input-octets] by rat-type*

... | *chart pie [(acct-output-octets + acct-input-octets) * 5] by rat-type*

... | *chart pie [acct-output-octets + acct-input-octets] [(acct-output-octets + acct-input-octets)/unique] by rat-type*

field_By..field_ByN Represents the aggregation dimension. It can be:

- one of the fields contained in sourcetype.
- one of the reserved words specified below:
  - *carrier*. This reserved word is used to aggregate by carrier and country.
  - *country*. This reserved word is used to aggregate by country.
  - *device*. This reserved word is used to aggregate by device type.
  - *uli-cell*. This reserved word is used to aggregate by user location.

-height=num is used to set a fixed height for the chart, where num is the number of pixels. If this parameter is not specified, default value will be used.

-width=num is used to set a fixed width for the chart, where num is the number of pixels. If this parameter is not specified, default value will be used.

-legend=no is used to hide the chart legend. This will apply only to the following charts: pie, col, bar, stackedbar, stackedcol and donut.

-ylabel=no is used to hide the y axis label.

-layout=tab|horizontal|vertical|gridWxH is used to decide how to represent the widgets when there are multiple metrics (in tabs or horizontally).

-override-title=yes is used to override the title of the panel with the information coming from the KPI definition (if defined).

-table=no is used to hide table containing chart's data.

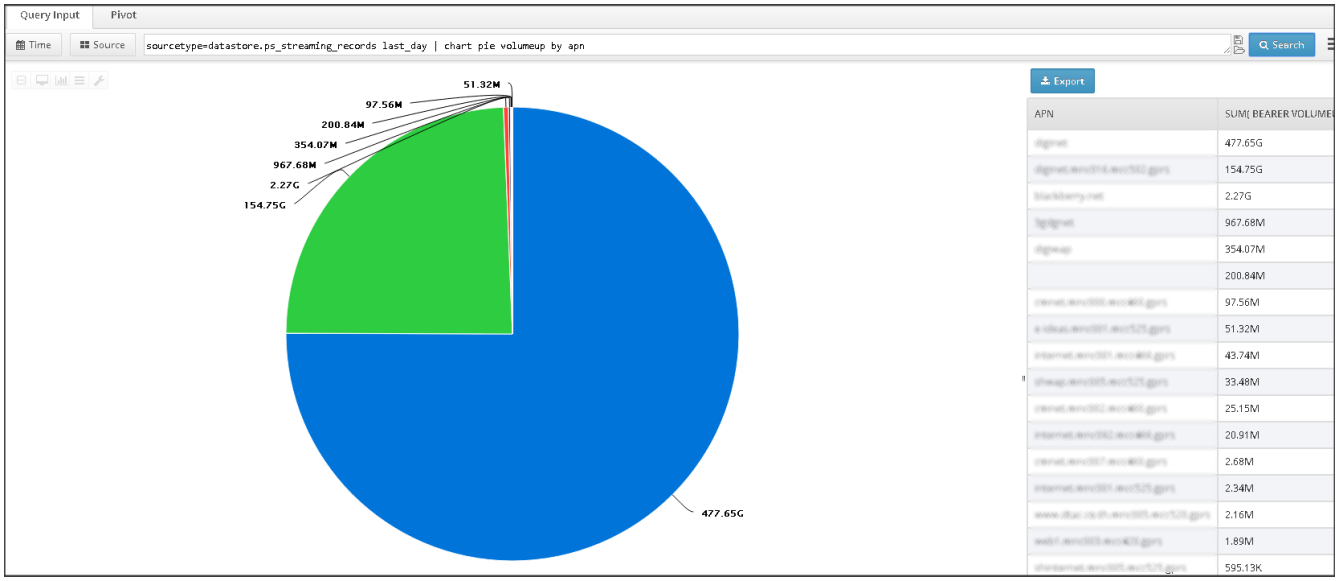
Example:

chart stackedbar3d volumedown volumeup by dpi-service -height=400 -width=400 -legend=no

Examples

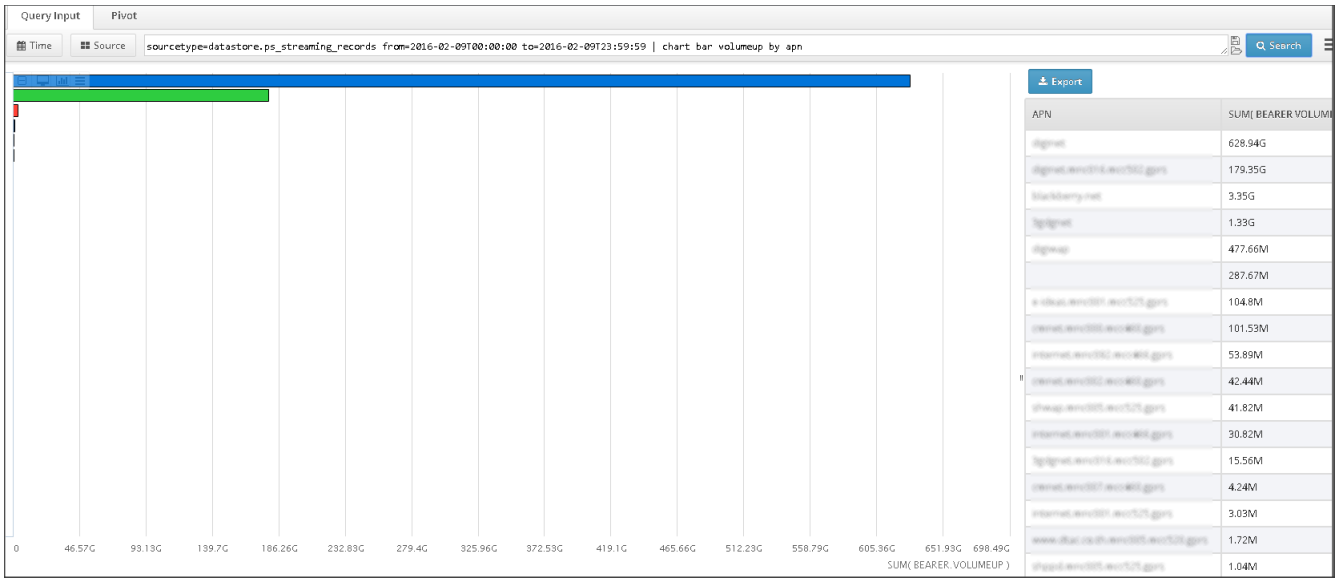
Example1. Pie chart showing volumeup of the streaming records aggregated by apn for a complete day.

sourcetype=datastore.ps_streaming_records last_day | chart pie volumeup by apn



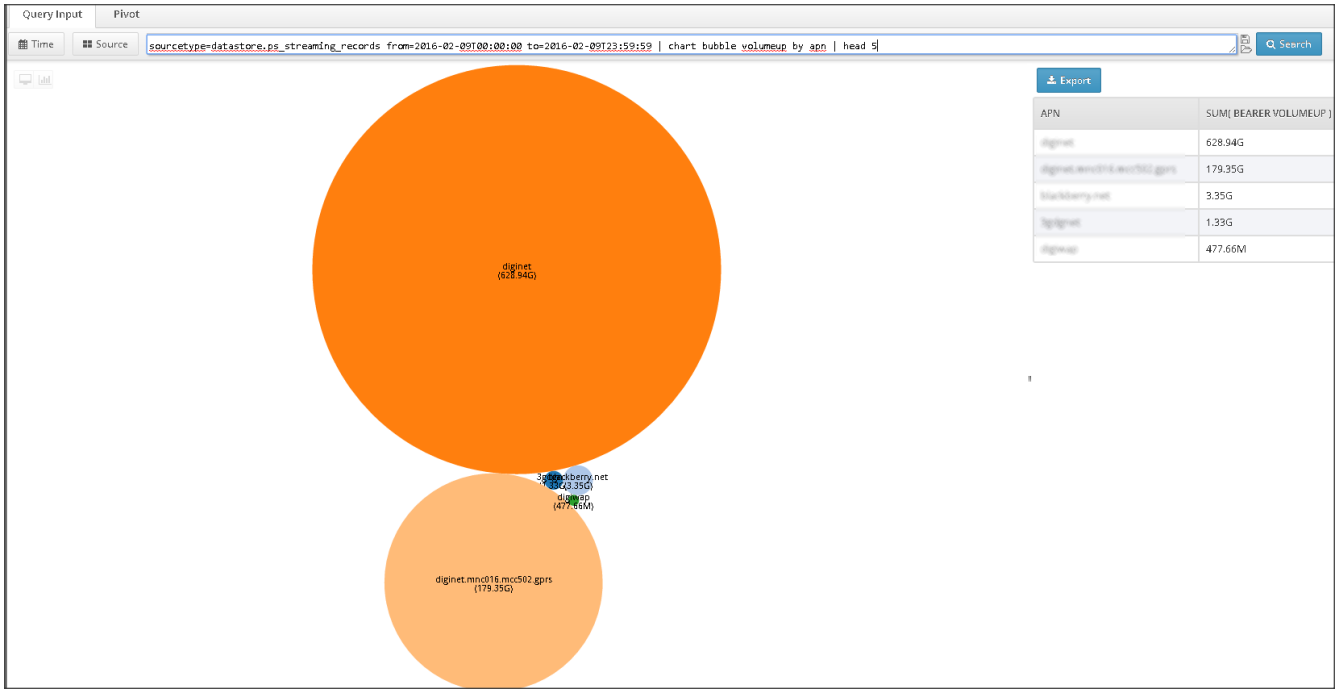
Example 2. Bar chart showing volume up of the streaming records aggregated by apn for a complete day.

sourcetype=datastore.ps_streaming_records from=2016-02-09T00:00:00 to=2016-02-09T23:59:59 | chart bar volumeup by apn



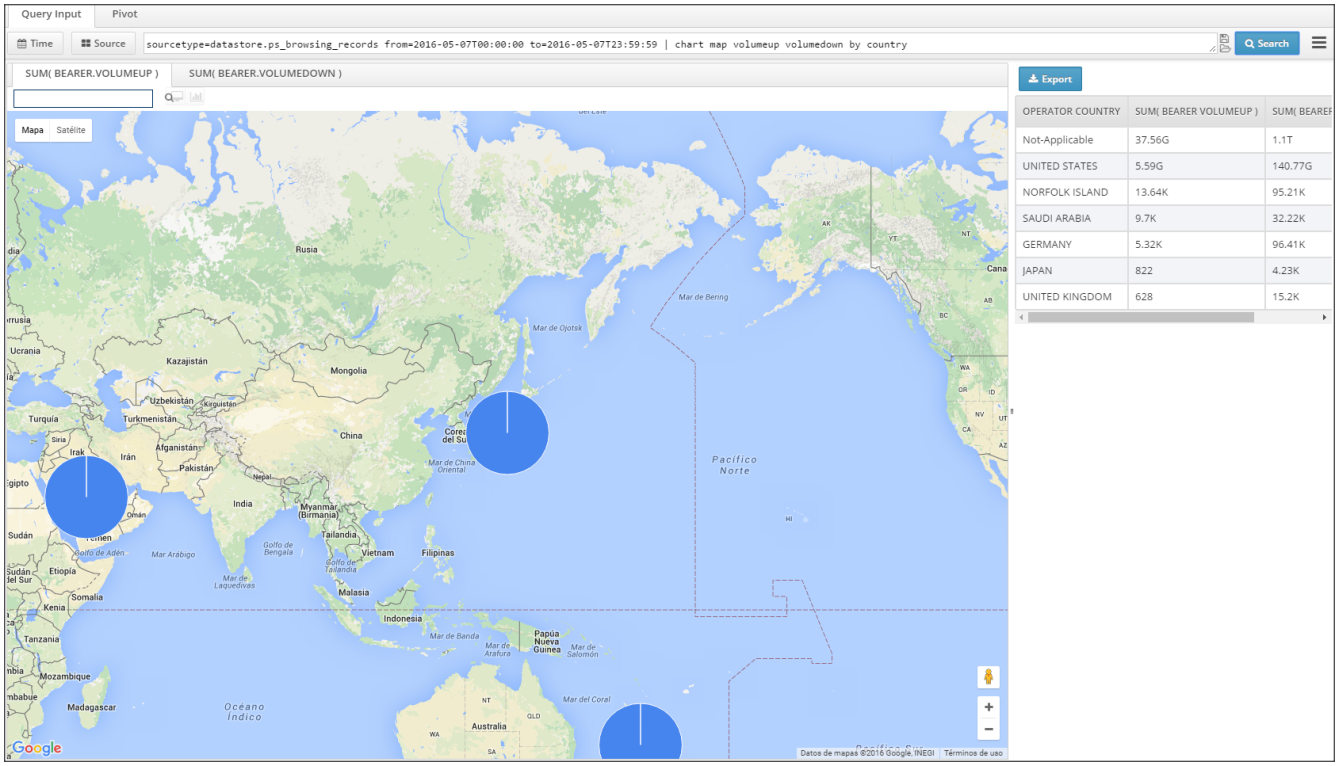
Example 3. Bar chart showing volume up of the streaming records aggregated by apn for a complete day.

sourcetype=datastore.ps_streaming_records from=2016-02-09T00:00:00 to=2016-02-09T23:59:59 | chart bubble volumeup by apn | head 5



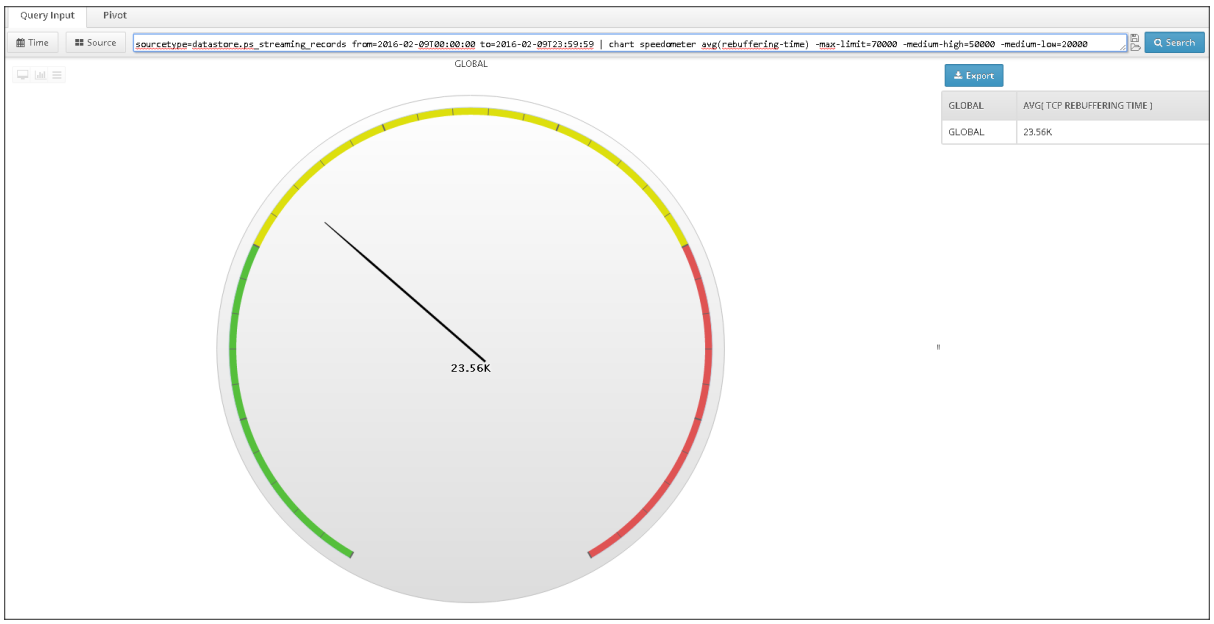
Example 4. Map chart showing volume up of the browsing records aggregated by country for a complete day.

sourcetype=datastore.ps_browsing_records from=2016-02-09T00:00:00 to=2016-02-09T23:59:59 | chart map volumeup volumedown by country



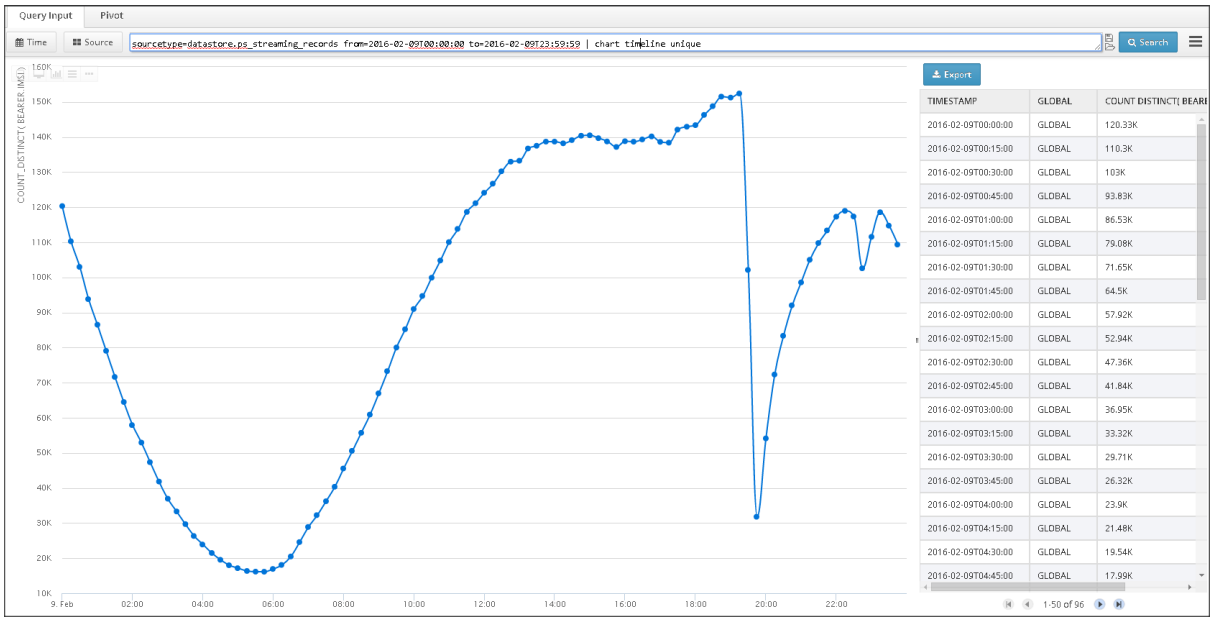
Example 5. Speedometer showing average rebuffering time for the complete network in a day.

sourcetype=datastore.ps_streaming_records from=2016-02-09T00:00:00 to=2016-02-09T23:59:59 | chart speedometer avg(rebuffering-time) -max-limit=70000 -medium-high=50000 -medium-low=20000



Example 6. Temporal evolution of unique subscribers with streaming records.

```
sourcetype=datastore.ps_streaming_records from=2016-02-09T00:00:00 to=2016-02-09T23:59:59 | chart timeline unique
```



3.6.2 cdf

The Cumulative Distribution Function (CDF) shows the probability that a given field fieldName takes on a value less than or equal to "x".

Syntax

cdf fieldName [by aggrField1 ...[aggrFieldN]] {params num-buckets min-value max-value} | {list val1 ...[valN]} [-unique=yes|no] [-max-results=num]

Parameters

fieldName: The field on which the cdf will be calculated

Example:

```
... | cdf acct-output-octets ...
```

by aggrField1 ...[aggrFieldN]: Aggregates the results by aggrField1...aggrFieldN

Example:

```
... | cdf acct-output-octets by APN ...
```

params num-buckets min-value max-value: Sets the minimum and maximum values and the number of buckets to calculate the probability distribution function.. Being:

- *num-buckets*. The number of buckets to be used .
- *min-value*. The minimum value.
- *max-value*. The maximum value.

Example:

```
... | cdf acct-output-octets parameters=3,0,9000
```

list val1...[valN]]: This parameter allows the buckets definition without the restriction of a uniform step value. Being:

val1...[valN] The buckets that are to be used.

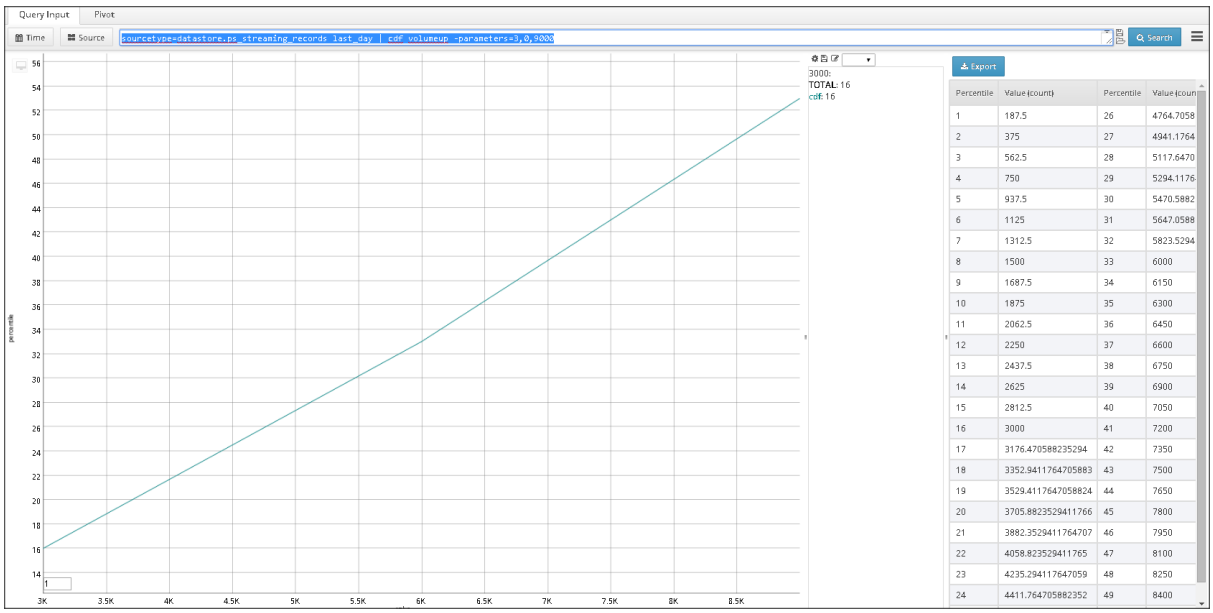
Example:

```
... | cdf acct-output-octets list 10 100 400 5000 8000 9000
```

-unique=yes | no: When set to 'yes', instead of take into account all the selected field's events, will take into account only the uniques selected fields's events. If this option is not present, it defaults to -unique=no

Example 1. Cumulative Distribution Function for the streaming volumeup

sourcetype=datastore.ps_streaming_records last_day | cdf volumeup -parameters=3,0,9000



3.6.3 pdf

The Probability Distribution Function (PDF) shows how the samples(%) are distributed among the different possible values for a given field fieldName.

Syntax

pdf fieldName [by aggrField1 ...[aggrFieldN]] {params num-buckets min-value max-value} | {list val1 ...[valN]} [-unique=yes|no] [-max-results=num]

Parameters

fieldName: The table's field on which the PDF will be calculated

Example:

```
... | pdf acct-output-octets ...
```

by aggrField1 ...[aggrFieldN]: Aggregates the results by aggrField1...aggrFieldN

Example:

```
... | pdf acct-output-octets by aggf1 aggf2 ...
```

params num-buckets min-value max-value: Sets the minimum and maximum values and the number of buckets to calculate the probability distribution function.. Being:

- *num-buckets*. The number of buckets to be used .
- *min-value*. The minimum value.
- *max-value*. The maximum value.

Example:

```
... | pdf acct-output-octets params 3 0 9000
```

list val1...[valN]: This parameter allows the buckets definition without the restriction of a uniform step value. Being:

- val1...[valN]. The buckets that are to be used.

Example:

```
... | pdf acct-output-octets list 10 100 400 5000 8000 9000
```

-unique=yes|no: When set to 'yes', instead of take into account all the selected field's events, will take into account only the uniques selected fields's events. If this option is not present, it defaults to -unique=no

Example:

```
... | pdf acct-output-octets params 3 0 9000 -unique=yes
```

-max-results=num: The max number of results (rows) that will be shown. By default num will be 200 events. Maximum value for num can be 100.000 events.

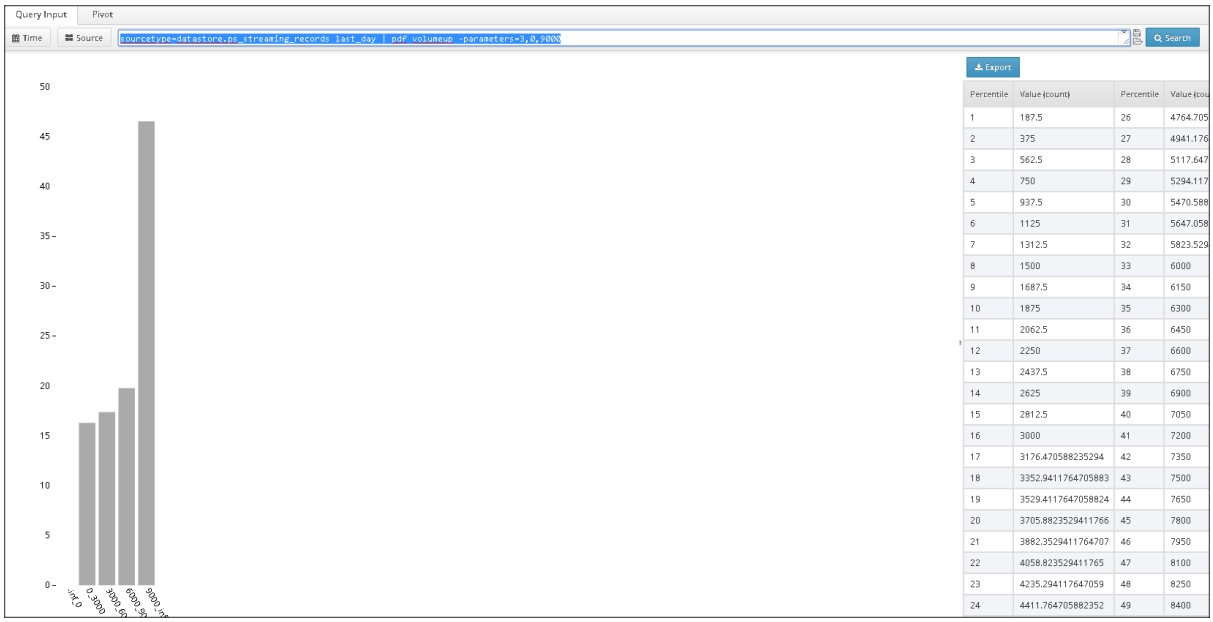
Example:

```
... | pdf acct-output-octets params 3 0 9000 -unique=yes -max-results=500
```

Example :

PDF for the streaming volumeup

sourcetype=datastore.ps_streaming_records last_day | pdf volumeup -parameters=3,0,9000



3.6.4 download

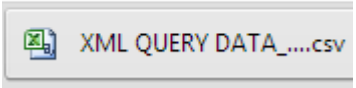
The download function executes the direct download of the query to an external file.

Syntax

download

Example 1. Exportation of the clustering of users as per their streaming service characterization based on the following indicators: tcp retransmissions, reproduction quality, rebuffering time and service access time.

```
from=2015-02-11T00:00:00 to=2015-02-11T23:59:59
sourcetype=datastore.pdu_streaming_records | pdf volumeup params 100 0 500000 |
download
```



3.6.5 filter

The filter function displays a table with the chosen fields. This function cannot be used with others function types.

Syntax

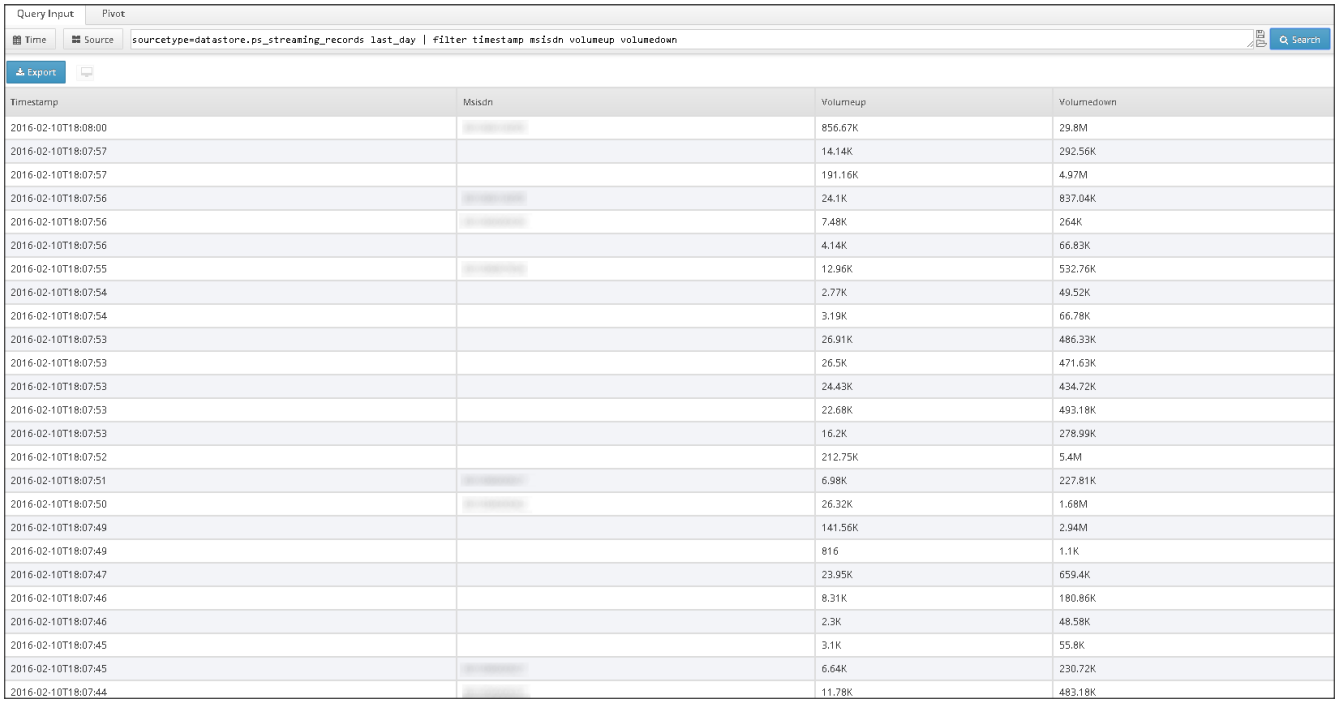
filter field1 ... [fieldN]

field1...fieldN are table fields of the specified sourcetype.

Examples:

```
... | filter acct-input-octets rat-type ras-client
```

sourcetype=datastore.ps_streaming_records last_day | filter timestamp msisdn volumeup volumedown



Timestamp	Msisdn	Volumeup	Volumedown
2016-02-10T18:08:00		856.67K	29.8M
2016-02-10T18:07:57		14.14K	292.56K
2016-02-10T18:07:57		191.16K	4.97M
2016-02-10T18:07:56		24.1K	837.04K
2016-02-10T18:07:56		7.48K	264K
2016-02-10T18:07:56		4.14K	66.83K
2016-02-10T18:07:55		12.96K	532.76K
2016-02-10T18:07:54		2.77K	49.52K
2016-02-10T18:07:54		3.19K	66.78K
2016-02-10T18:07:53		26.91K	486.33K
2016-02-10T18:07:53		26.5K	471.63K
2016-02-10T18:07:53		24.43K	434.72K
2016-02-10T18:07:53		22.68K	493.18K
2016-02-10T18:07:53		16.2K	278.99K
2016-02-10T18:07:52		212.75K	5.4M
2016-02-10T18:07:51		6.98K	227.81K
2016-02-10T18:07:50		26.32K	1.68M
2016-02-10T18:07:49		141.56K	2.94M
2016-02-10T18:07:49		816	1.1K
2016-02-10T18:07:47		23.95K	659.4K
2016-02-10T18:07:46		8.31K	180.86K
2016-02-10T18:07:46		2.3K	48.58K
2016-02-10T18:07:45		3.1K	55.8K
2016-02-10T18:07:45		6.64K	230.72K
2016-02-10T18:07:44		11.78K	483.18K

3.6.6 get-sessions

The get sessions function allows obtaining a session summary from a signaling based source. It lookup the newest reference for a session, which is denoted by the subscriber and session identifiers.

Syntax

get-sessions -subscriber-field=subscriberField -session-field=sessionField [-jitter=jitter -offset-seek=offsetSeek -default-start-time-field=timestampField -default-end-time-field=timestampField]

Parameters

-subcriber-field: Name of the subcriber id field.

Example:

```
-subscriber-field=bearer.msisdn
```

-session-field: Name of the session id field. It can also be specified as a list of fields

Example:

```
-session-field=gtp.session-id
```

Example:

```
-session-field=gtp.field1,gtp.field2
```

-*default-start-time-field*: Field to consider for start session time. Set this if differs the event timestamp

Example:

```
-default-start-time-field=event.start-time
```

-*default-end-time-field:* Field to consider for end session time. Set this if differs the event timestamp

Example:

```
-default-end-time-field=event.stop-time
```

-*jitter:* Session jitter in seconds. Will be included in session start and end timestamps

Example:

```
-jitter=3600
```

-*offset-seek:* offset in seconds to considerer sessions messages which are not within the query from and to

Example

-offset-seek=7200

-*promoted-fields*: fields to show in the result. If any provided field doesn't exist, it will be discarded

Example

-promoted-fields=msisdn,imsi,apn

Example

from=2015-01-22T00:00:00 to=2015-01-22T23:59:59 sourcetype=datastore.ps_session_records <msisdn>| get-sessions -subscriber-field=bearer.msisdn -session-field=gtp.session.id

3.6.7 head

The head function limits the number of returned values in the performed query.

Syntax

head rows_number

rows_number is a number. This is the number of rows to be returned.

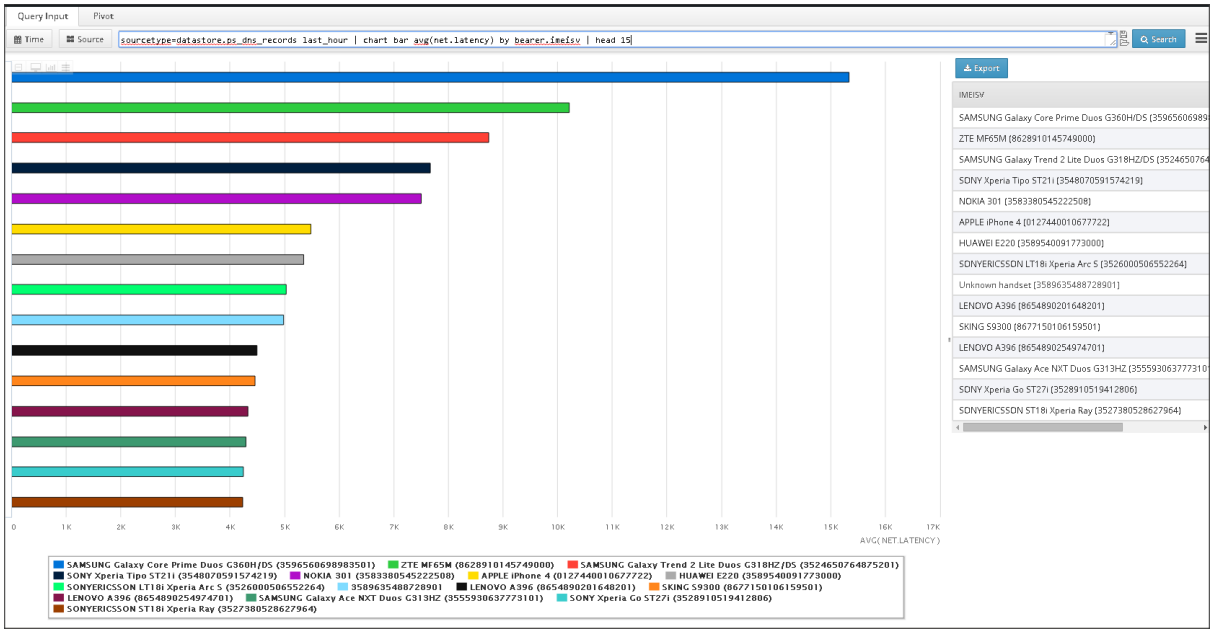
Examples:

... | head 2

... | filter msisdn rat-type | head 2

Query *Example* 1: top 15 devices per average latency

sourcetype=datastore.ps_dns_records last_hour | chart bar avg(net.latency) by bearer.imeisv | head 15



3.6.8 heatmap

The heatmap function displays a heat map based on the type chosen through the different available fields in the records.

Heat maps provide a graphical representation of the metrics in a map. The metric values are clustered by colors in different layers.

Syntax

heatmap [field [fieldN]] [order-by orderField [orderFieldN] [asc|desc]] [by groupbyField [groupbyFieldN]] [-table=yes] [-results=num] [-gradient-type=num] [-single-layer=yes] [-format-results=yes]

Parameters

field..fieldN can be one of the following elements:

- One of the table fields specified in sourcetype.
- A function:
  - avg(fieldA). This function calculates the average value of the field specified in fieldA.
  - median(fieldA). This function calculates the median value of the field specified in fieldA.
  - max(fieldA). This function calculates the maximun value of the field specified in fieldA.
  - min(fieldA). This function calculates the minimun value of the field specified in fieldA.
  - sum(fieldA). This function sum the values of the field specified in fieldA.
  - count(fieldA). This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified.
  - unique(fieldA). This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified

[fieldA operator fieldB] [as expressionName]: an expression or group of expressions in brackets. Where:

fieldA, fieldB can be:

- one of the table fields specified in sourcetype.
- numerical values.
- a function:
  - avg(fieldA). This function calculates the average value of the field specified in fieldA.
  - median(fieldA). This function calculates the median value of the field specified in fieldA.
  - max(fieldA). This function calculates the maximun value of the field specified in fieldA.
  - min(fieldA). This function calculates the minimun value of the field specified in fieldA.
  - sum(fieldA). This function sum the values of the field specified in fieldA.
  - count(fieldA). This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified.
  - unique(fieldA). This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified

.

- *operator* can be any of the arithmetic operators: +, -, *, /
- *as expressionName* to specify an alias for the expression. This parameter is optional.

order-by orderField..orderFieldN [asc|desc]:

to order results by an specific criteria. Being:

- *orderField..orderFieldN* must be one of the fields specified in the heatmap fields mentioned above.
- *asc | desc:* to specify if you want ascendant or descendant order. Optional. By default, order is ascendant.

by groupByField..groupByFieldN: to create a multilayer heatmap by group-by fields.

-table=yes is used to show the result table. By default, the legend is hidden.

-results=num is the max number of results shown in the map. By default num will be 10.000 events. Maximum value fornum can be 100.000 events. It can be used with "max" value.

-gradient-type=num is the type of colors used to represent the heatmap. By default num will be green to red. Color values are:

0 = red

1 = green

2 = blue

3 = yellow

4 = magenta

5 = cyan

-single-layer=yes this forces single-layer although groupByField(s) are present. By defauld, single-layer is only active if no groupByField(s) are present.

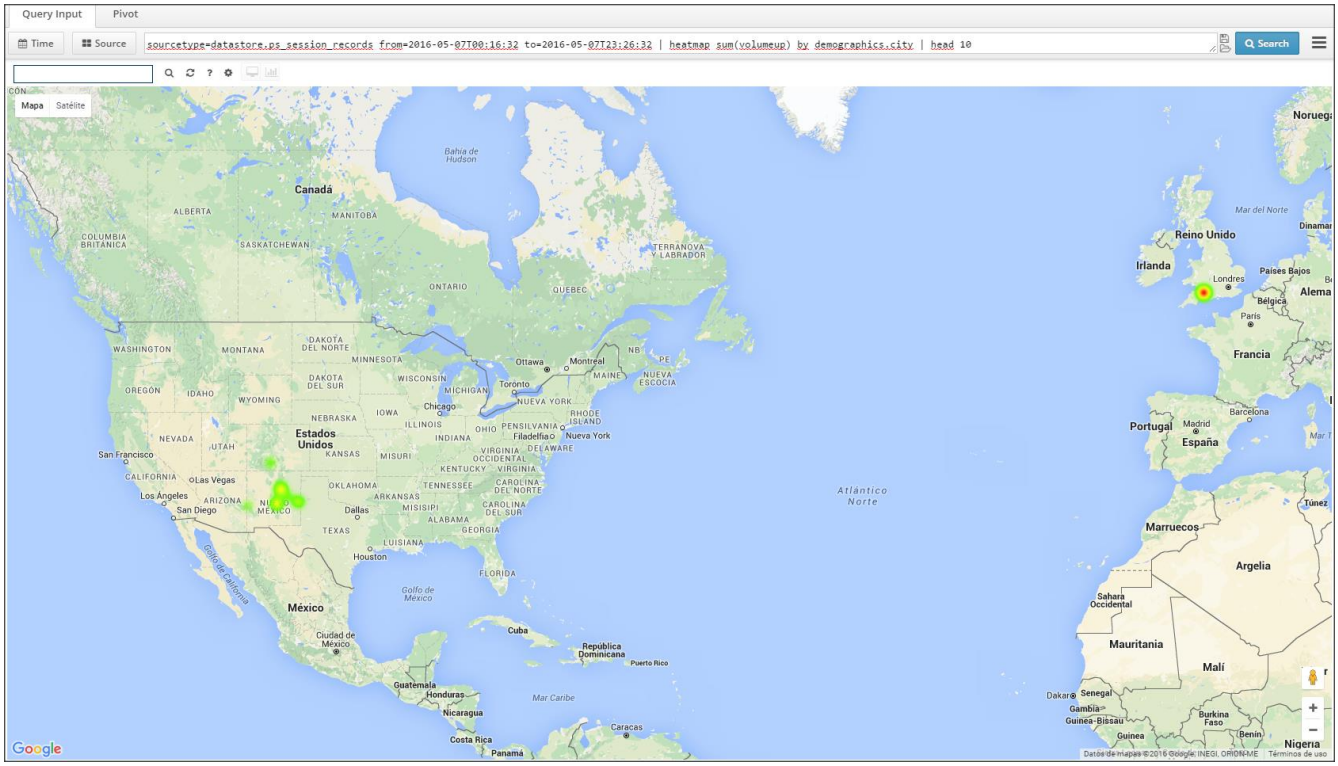
-format-results=yes this forces the result table results to be formatted (ie: Volume in K, M, G or T). Disabled by default.

Examples:

... | heatmap [AVG(acct-input-octets)] as avgVolume order-by avgVolume -format-results=yes

... | heatmap unique -table=yes order-by unique desc -results=max

sourcetype=datastore.ps_session_records from=2016-02-10T17:16:32 to=2016-02-10T17:26:32 | heatmap sum(volumeup) by rat-type-str uli-cell | head 10



3.6.9 kqi-chart

The kqi-chart is the function that allows the use of the precomputed reports in the Full Search Panel as part of a query.

Syntax

kqi-chart type "path_to_the_kqi_report" ["path_to_other_kqi_report"] by field_By [field_ByN] [-height=num] [-width=num] [-legend=no] [-table=no] [-ylabel=no] [-layout=tab|horizontal|vertical|gridWxH] [-include-units=no] [-override-title=no] [-include-thresholds=no] [-table-position=north|south|east|west]*

Parameters

type can be one of the elements contained in Table 2

"path_to_the_kqi_report" is the path where the kqi report is stored. The available kqi report paths are automatically visible and selectable in the on-line help once you have selected the data source and the kqi-chart operator.

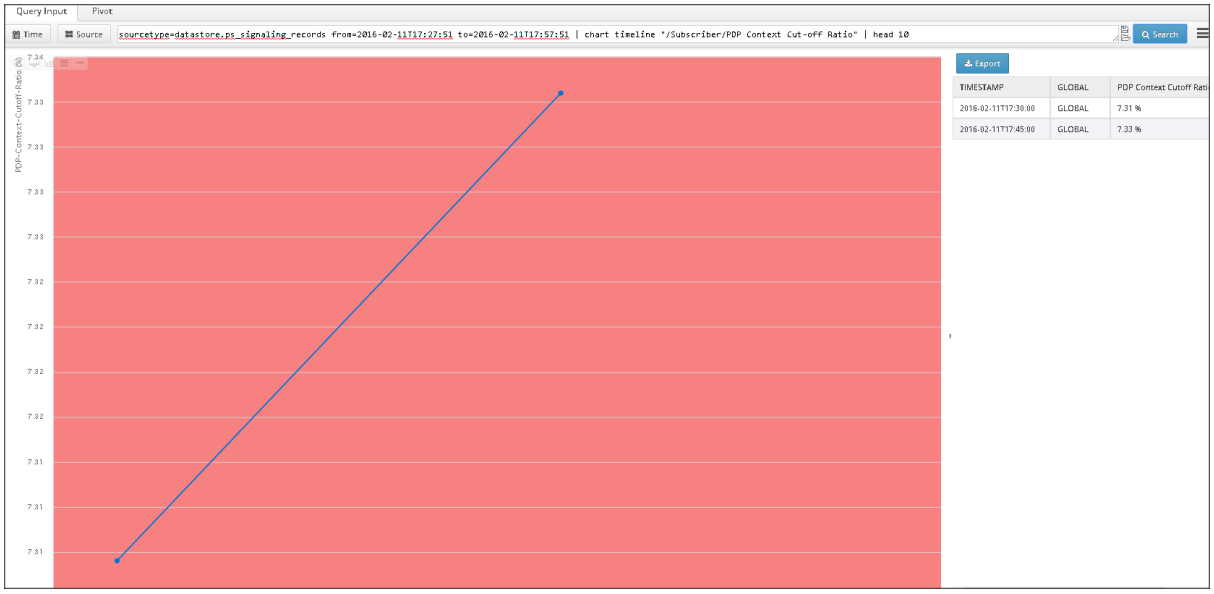
- *field_By..field_ByN* can be:
  - one of the table fields specified in sourcetype.
  - one of the reserved words specified below:
    - carrier. This reserved word is used to aggregate by carrier and country.
    - country. This reserved word is used to aggregate by country.
    - device. This reserved word is used to aggregate by device type.
    - uli-cell. This reserved word is used to aggregate by user location.
- *-height=num is* used to set a fixed height for the chart, where num is the number of pixels. If this parameter is not specified, default value will be used.
- *-width=num* is used to set a fixed width for the chart, where num is the number of pixels. If this parameter is not specified, default value will be used.

- *-legend=no* is used to hide the chart legend. This will apply only to the following charts: pie, col, bar, stackedbar, stackedcol and donut.
- *-table=no* is used to hide table containing kqi-chart's data.
- *-ylabel=no* is used to hide the y axis label.
- *-layout=tab|horizontal|vertical|gridWxH is used to decide how to represent the widgets when there are multiple metrics (in* tabs or horizontally).
- *-override-title=yes* is used to override the title of the panel with the information coming from the KPI definition (if defined).
- *-include-units=yes* is used include units in all tooltips.
- *-include-thresholds=yes* is used include thresholds in the widgets. Defaults to true.
- *-table-position=north|south|east|west* is used to include the table in one of the given positions when it is displayed. Defaults to east.

Example:

kqi-chart pie "/Axis-Network/GGSN/Total throughput" by ggsn -height=400 -width=400 -legend=no

sourcetype=datastore.ps_signaling_records from=2016-02-11T17:27:51 to=2016-02-11T17:57:51 | chart timeline "/Subscriber/PDP Context Cut-off Ratio" | head 10



sourcetype=sourcetype=datastore.ps_signaling_records from=2016-02-11T17:27:51 to=2016-02-11T17:57:51 | chart timeline "/Subscriber/PDP Context Cut-off Ratio" by APN -table=no | head 10

Per subscribers available indicators

/Subscriber/Bearer Creation Failure Ratio

/Subscriber/Bearer Creation Failure Ratio (Global)

/Subscriber/Bearer Modification Failure Ratio

/Subscriber/DNS Host Name Resolution Failure Ratio

/Subscriber/DNS Host Name Resolution Time

/Subscriber/End-to-end Latency

/Subscriber/File Download and Upload Data Transfer Cut-off

/Subscriber/File Download and Upload Failure Ratio

/Subscriber/File Download and Upload IP Service Access Failure Ratio

/Subscriber/File Download and Upload IP Service Setup Time

/Subscriber/File Download and Upload Service Non-Accessibility

/Subscriber/File Download and Upload Setup Time

/Subscriber/File Download Effective Data Rate

/Subscriber/File Download Mean Data Rate

/Subscriber/File Download Session Time

/Subscriber/File Upload Effective Data Rate

/Subscriber/File Upload Mean Data Rate

/Subscriber/File Upload Session Time

/Subscriber/HTTP Mean Data Rate

/Subscriber/HTTP Peak Data Rate

/Subscriber/HTTP Session Failure Ratio

/Subscriber/HTTP Session Time

/Subscriber/Number of Videos

/Subscriber/PDP Context Creation Failure Ratio (Gn)

/Subscriber/PDP Context Creation Time (Gn)

/Subscriber/PDP Context Cut-off Ratio

/Subscriber/Session Creation Failure Ratio

/Subscriber/Session Creation Time

/Subscriber/Streaming Rebuffering Time

/Subscriber/Streaming Rebuffering Time Percentage

/Subscriber/Streaming Reproduction Quality

/Subscriber/TCP Retransmission Ratio

/Subscriber/TCP Round Trip Time (Client side)

/Subscriber/TCP Round Trip Time (Server side)

/Subscriber/Throughput as provided by the network (Mean Throughput)

/Subscriber/Throughput as provided by the network (Peak Throughput)

/Subscriber/Time to Stream Start

/Subscriber/Total Reproduction Time per Video

/Subscriber/Total Throughput

/Subscriber/Total Volume

/Subscriber/Total Volume (uplink+downlink)

/Subscriber/Video Stall per Video

3.6.10 kqi-heatmap

The kqi-heatmap is the function that allows the use of the precomputed reports as part of a heat map.

Syntax

kqi-heatmap ["path_to_the_kqi_report" ["path_to_other_kqi_report"]] [order-by orderField [orderFieldN] [asc|desc]] [by groupbyField [groupbyFieldN]] [-table=yes] [-results=num] [-gradient-type=num] [-single-layer=yes] [-format-results=yes]*

Parameters

order-by orderField..orderFieldN [asc|desc]: to order results by an specific criteria. Being:

- *orderField..orderFieldN* must be one of the fields specified in the heatmap fields mentioned above.
- *asc|desc:* to specify if you want ascendant or descendant order. Optional. By default, order is ascendant.

by groupByField..groupByFieldN: to create a multilayer heatmap by group-by fields.

-table=yes is used to show the result table. By default, the legend is hidden.

-results=num is the max number of results shown in the map. By default num will be 10.000 events. Maximum value fornum can be 100.000 events. It can be used with "max" value.

-gradient-type=num is the type of colors used to represent the heatmap. By default num will be green to red. Color values are:

- 0 = red
- 1 = green
- 2 = blue
- 3 = yellow
- 4 = magenta
- 5 = cyan

-single-layer=yes this forces single-layer although groupByField(s) are present. By defauld, single-layer is only active if no groupByField(s) are present.

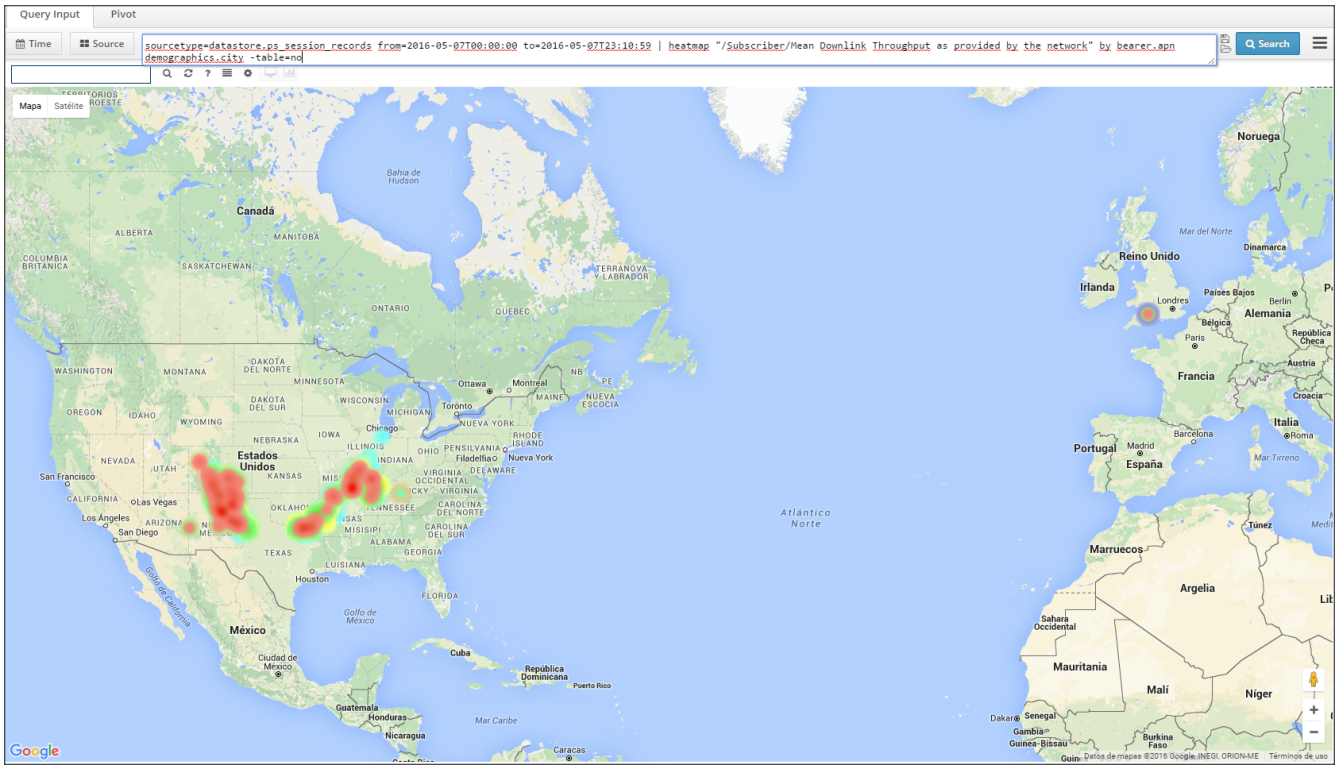
-format-results=yes this forces the result table results to be formatted (ie: Volume in K, M, G or T). Disabled by default.

Example

... | kqi-heatmap "/Axis-Location/City/Total throughput" -format-results=yes

Example

sourcetype=datastore.ps_session_records from=2016-05-07T00:00:00 to=2016-05-07T23:10:59 | heatmap "/Subscriber/Mean Downlink Throughput as provided by the network" by bearer.apn demographics.city -table=no



3.6.11 ladder

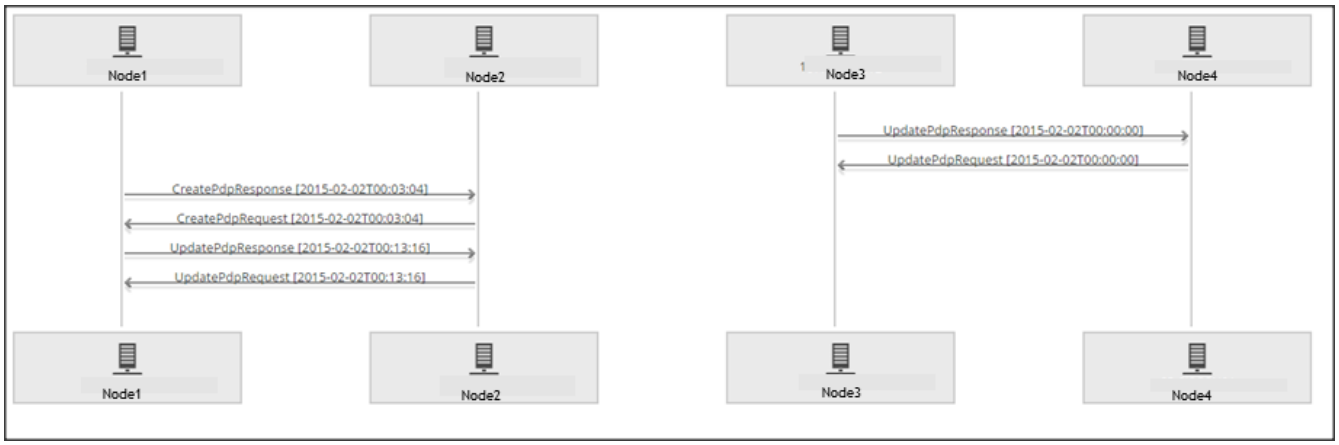
The ladder function provides a graphical representation of a ladder diagram by using the messages present in the records. It is aimed to troubleshooting representation of the signaling messages

Syntax

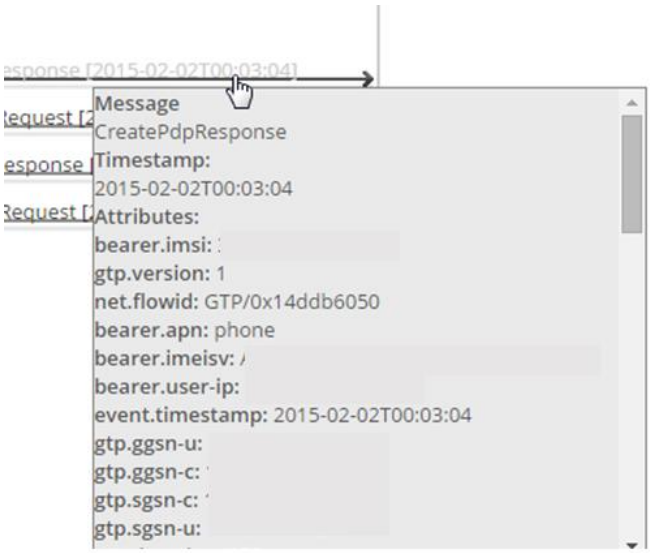
Ladder

Examples

from=2016-02-02T00:00:00 to=2016-02-02T00:14:59 sourcetype=datastore.ps_signaling_records <msisdn> | ladder



Note that ladder may be interesting for a specific user. Each arrow represents a signaling message from one node to another node. By over hovering each arrow the details of the message will pop up as shown in the next figure:



3.6.12 ladder-tree

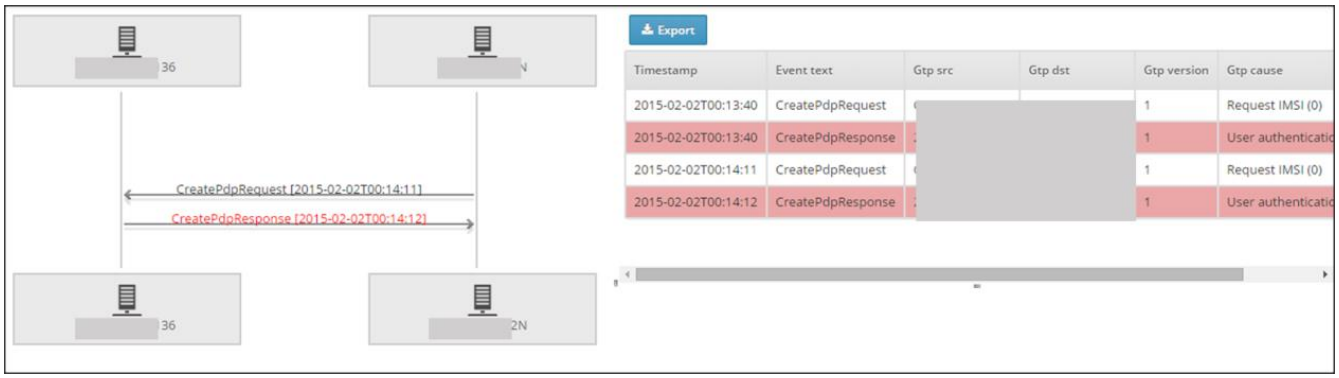
The ladder-tree function will provide a graphical representation of a ladder diagram and a summarize table by using the messages present in the records. By each table entry, after click, a tree with transaction details will be shown

Syntax

ladder-tree

Example

from=2016-02-02T00:00:00 to=2016-02-02T00:14:59 sourcetype=datastore.ps_signaling_records xxxxx | ladder-tree



3.6.13 map

The map function provides a geographical representation of the location information in the records.

Syntax

map [by uli-cell] [-table-position=north|south|east|west]

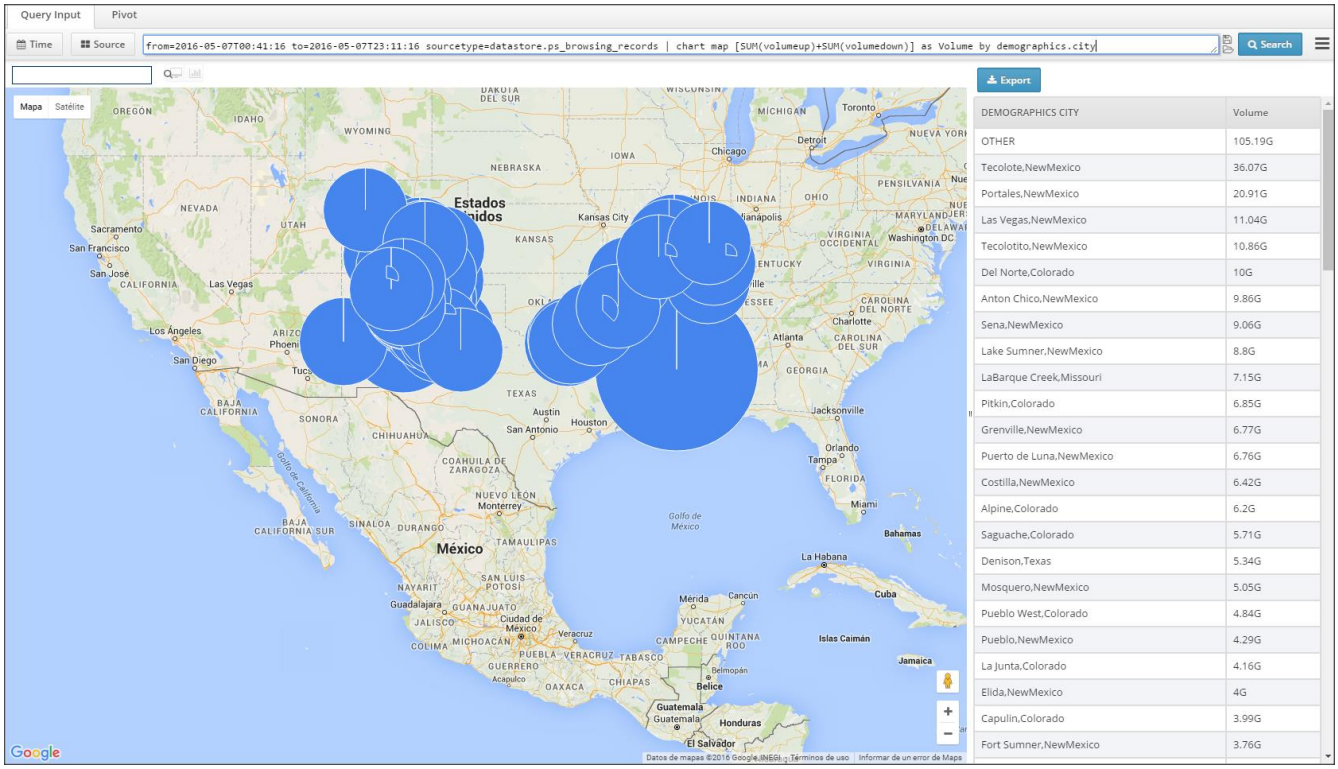
Parameters

by uli-cell: to use cell info in order to geolocate the event.

-table-position=north|south|east|west is used to include the table in one of the given positions when it is displayed. Defaults to east.

Example:

- *from=2016-02-11T17:41:16 to=2016-02-11T18:11:16 sourcetype=datastore.ps_browsing_records | chart map [SUM(volumeup)+SUM(volumedown)] as Volume by uli-cell*



3.6.14 multi-source

The Multi Source function allows to show information from different data sources when they have a common field with a specific value.

This function can be used for *example* to see all the information for a specific MSISDN in different data sources.

Syntax

> *multi-source table2 ...[tableN] -fieldname=fieldTable1,...,fieldTableN -fieldvalue=value [-columnfield=column1,...,columnN] [-charttype=table|ladder]*

Parameters

table2 ...[tableN]: The tables where the query will be performed sourcetype table in addition to the one specified in the sourcetype option. Clarification: table1 would be the table defined in the sourcetype option.

Example:

... sourcetype = datastore.radius | multi-source datastore.mme ...

-fieldname=fieldTable1,...,fieldTableN: The name of the field to be filtered For each corresponding table. They can have different names but must represent the same field. fieldTable1 is the field associated to the table defined in sourcetype option, fieldTable2 is the field associated to table2 and so on...

Example:

... sourcetype = datastore.radius | multi-source datastore.mme -fieldname=calling-station-id,msisdn ...

-fieldvalue=value: The value to be filtered for the defined tables

Example:

sourcetype=datastore.radius | multi-source datastore.mme -fieldname=calling-station-id,msisdn -fieldvalue=34651xxxxxx

-columnfield=column1,...,columnN: Optional. A comma separated list of additional columns(of any of the selected tables) to be shown in the multi-source table when charttype=table. Note that configuring this field any other configuration will be discarded

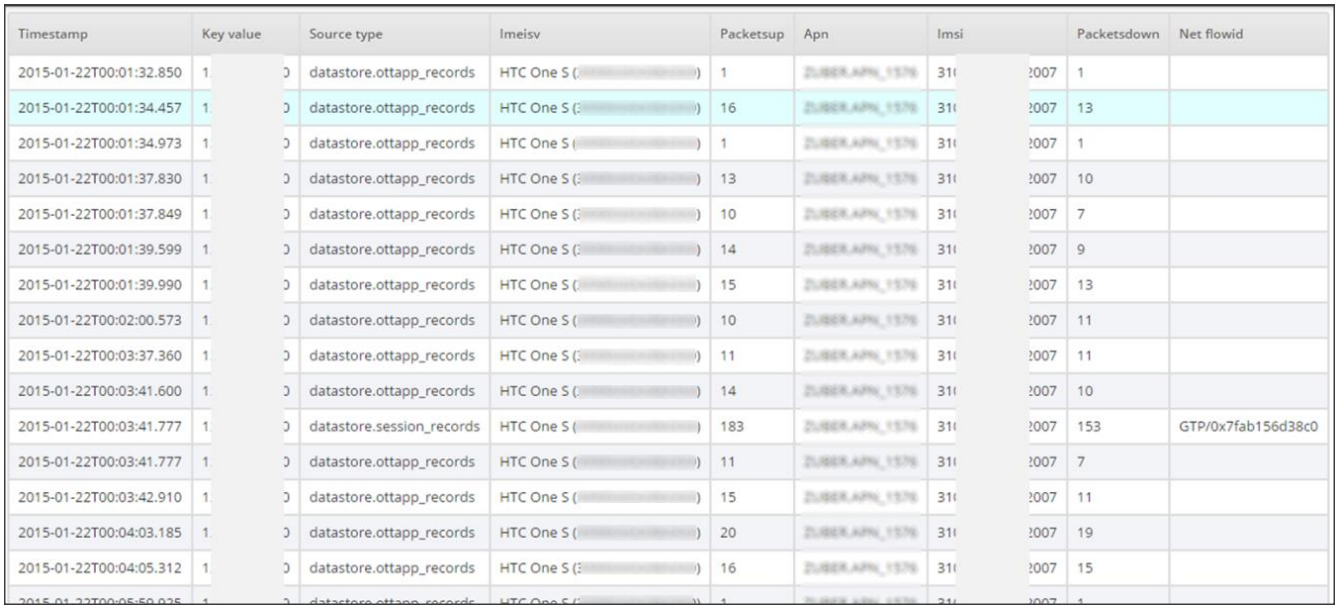
Example:

sourcetype=datastore.radius | multi-source datastore.pcrf -fieldname=calling-station-id,msisdn -fieldvalue=34651xxxxxx -columnfield=acct-output-octets,framed-ip-address,full-name,acct-session-time,id

-charttype=table|ladder: Optional. When set to table it will show a table with the results of the query. When set to ladder it will show a ladder + table with the results of the query. Ladder will only work when gui hint config files are properly configured. When this option is not present it defaults to -charttype=table

Example:

from=2015-01-22T00:00:00 to=2015-01-22T23:59:59 sourcetype=datastore.session_records | multi-source datastore.pdu_dns_records -fieldname=msisdn,msisdn -fieldvalue=<msisdn>-charttype=table



Timestamp	Key value	Source type	Imeisv	Packetsup	Apn	Imsi	Packetsdown	Net flowid
2015-01-22T00:01:32.850	1	datastore.ottapp_records	HTC One S (	1			1	
2015-01-22T00:01:34.457	1	datastore.ottapp_records	HTC One S (	16			13	
2015-01-22T00:01:34.973	1	datastore.ottapp_records	HTC One S (	1			1	
2015-01-22T00:01:37.830	1	datastore.ottapp_records	HTC One S (	13			10	
2015-01-22T00:01:37.849	1	datastore.ottapp_records	HTC One S (	10			7	
2015-01-22T00:01:39.599	1	datastore.ottapp_records	HTC One S (	14			9	
2015-01-22T00:01:39.990	1	datastore.ottapp_records	HTC One S (	15			13	
2015-01-22T00:02:00.573	1	datastore.ottapp_records	HTC One S (	10			11	
2015-01-22T00:03:37.360	1	datastore.ottapp_records	HTC One S (	11			11	
2015-01-22T00:03:41.600	1	datastore.ottapp_records	HTC One S (	14			10	
2015-01-22T00:03:41.777	1	datastore.session_records	HTC One S (	183			153	GTP/0x7fab156d38c0
2015-01-22T00:03:41.777	1	datastore.ottapp_records	HTC One S (	11			7	
2015-01-22T00:03:42.910	1	datastore.ottapp_records	HTC One S (	15			11	
2015-01-22T00:04:03.185	1	datastore.ottapp_records	HTC One S (	20			19	
2015-01-22T00:04:05.312	1	datastore.ottapp_records	HTC One S (	16			15	

3.6.15 network

The network function will display the network by auto-discovering it through the different available fields in the records.

Syntax

network type [-path=path] field [fieldN] [order-by orderField [orderFieldN] [asc|desc]]

Parameters

type can be one of the following elements:

- *rgraph*. Displays the network as an RGraph. This is the default representation if no type is specified in the command line.
- *hypertree*. Displays the network as a hypertree.
- *treemap*. Displays the network as a treemap.
- *spacetree*. Displays the network as a tree.

Example:

... | *network treemap*

-path=pathFields can be used to show the nodes and connections in network representation.

- *pathFields* can be one of the table fields specified in sourcetype separated by ">" with no blank spaces.

Example: Representation of the volume transferred between sgsn and ggsn nodes.

... | *network -path=3gpp-sgsn-address>ras-client SUM(acct-input-octets)*

field..fieldN can be one of the following elements:

- one of the table fields specified in sourcetype.
- a function:
  - avg(fieldA). This function calculates the average value of the field specified in fieldA.
  - median(fieldA). This function calculates the median value of the field specified in fieldA.
  - max(fieldA). This function calculates the maximun value of the field specified in fieldA.
  - min(fieldA). This function calculates the minimun value of the field specified in fieldA.
  - sum(fieldA). This function sum the values of the field specified in fieldA.
  - count(fieldA). This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified.
  - unique(fieldA). This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified.
  - avg_if(fieldA, expression). This function calculates the average value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
  - median_if(fieldA, expression). This function calculates the median value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
  - max_if(fieldA, expression). This function calculates the maximun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
  - min_if(fieldA, expression). This function calculates the minimun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
  - sum_if(fieldA, expression). This function sum the values of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
  - count_if(fieldA, expression). This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified when the expression defined is true. Available expressions are the same as the ones used in where function

- unique_if(fieldA, expression). This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified when the expression defined is true. Available expressions are the same as the ones used in where function

- *[fieldA operator fieldB] [as expressionName] :* an expression or group of expressions in brackets.
  - *fieldA, fieldB can* be one of the following elements:
    - one of the table fields specified in sourcetype.
    - numerical values.
    - a function:
      - avg(fieldA). This function calculates the average value of the field specified in fieldA.
      - median(fieldA). This function calculates the median value of the field specified in fieldA.
      - max(fieldA). This function calculates the maximun value of the field specified in fieldA.
      - min(fieldA). This function calculates the minimun value of the field specified in fieldA.
      - sum(fieldA). This function sum the values of the field specified in fieldA.
      - count(fieldA). This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified.
      - unique(fieldA). This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified.
      - avg_if(fieldA, expression). This function calculates the average value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
      - median_if(fieldA, expression). This function calculates the median value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
      - max_if(fieldA, expression). This function calculates the maximun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
      - min_if(fieldA, expression). This function calculates the minimun value of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
      - sum_if(fieldA, expression). This function sum the values of the field specified in fieldA when the expression defined is true. Available expressions are the same as the ones used in where function
      - count_if(fieldA, expression). This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified when the expression defined is true. Available expressions are the same as the ones used in where function
      - unique_if(fieldA, expression). This function counts the unique occurrences number of the field specified in fieldA or the primary key

if no fieldA is specified when the expression defined is true. Available expressions are the same as the ones used in where function

- *operator* can be any of the arithmetic operators: +, -, *, /
- *as expressionName:* to specify an alias for the expression. This parameter is optional.

Examples:

... | network [avg(acct-output-octets) + avg(acct-input-octets)] as avgVol order-by avgVol desc

... | network [avg(acct-output-octets) / 3] as avgVolThird order-by avgVolThird desc

order-by orderField..orderFieldN [asc | desc]: to order results by an specific criteria. Being:

- orderField..orderFieldN must be one of the fields specified in the network fields mentioned above.
- asc|desc: to specify if you want ascendant or descendant order. Optional. By default, order is ascendant.

Example:

*... | network [AVG(acct-input-octets) * 4] as avg4Times order-by avg4Times*

3.6.16 pivot-table

The pivot-table function will display a pivot table with the rows and columns passed as group by **parameters**.

Syntax

pivot-table field [fieldN] by field_Row_By [field_Row_ByN] and field_Column_By [field_Column_ByN]

Parameters

field [fieldN] can be one of the following elements:

- One of the fields contained in the query sourcetype.
- A function included in Table 3.

Table 5. <Function> Parameter Description

<function>	Meaning
avg(fieldA)	This function calculates the average value of the field specified in fieldA. Where fieldA is one of fields contained in the sourcetype.
median(fieldA)	This function calculates the median value of the field specified in fieldA
max(fieldA)	This function calculates the maximun value of the field specified in fieldA
min(fieldA)	This function calculates the minimun value of the field specified in fieldA
sum(fieldA)	This function sum the values of the field specified in fieldA

count(fieldA)	This function count the number of events of the field specified in fieldA or the primary key if no fieldA is specified
unique(fieldA)	This function counts the unique occurrences number of the field specified in fieldA or the primary key if no fieldA is specified

- *[fieldA operator fieldB] [as expressionName]* which is an expression or group of expressions in brackets, where:
  - *fieldA, fieldB* can be:
    - one of the fields included in sourcetype.
    - numerical values
    - a function as described in Table 3.
  - *operato*r can be any of the following arithmetic operators included in Table 4
  - *as expressionName* it is an optional parameter to specify an alias for the expression

Table 6. <operator> Parameter Description

<operator>	**Meaning**
+	Addition
-	Subtraction
*	Multiplication
/	Division

Examples

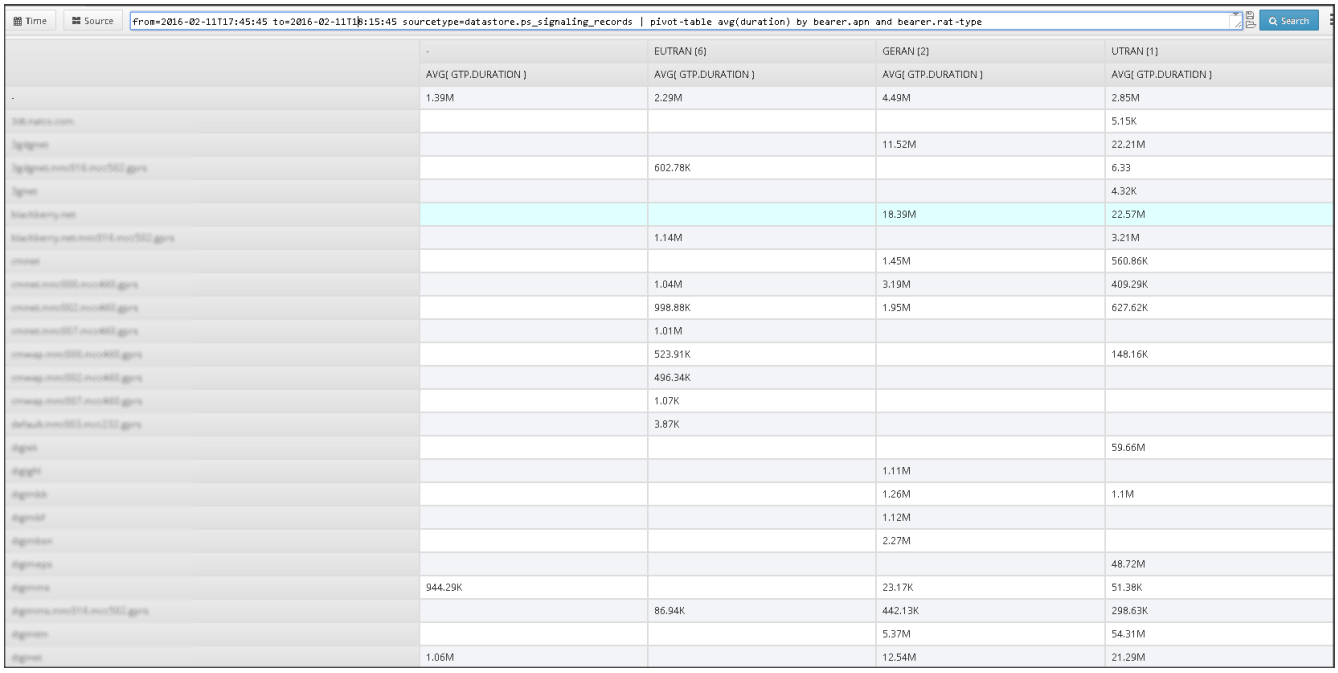
... | pivot-table [acct-output-octets + acct-input-octets] as volume by country and rat-type

... | pivot-table [acct-output-octets + acct-input-octets] by rat-type

*... | pivot-table [(acct-output-octets + acct-input-octets) * 5] by rat-type*

... | pivot-table [acct-output-octets + acct-input-octets] [(acct-output-octets + acct-input-octets)/unique] by rat-type

from=2016-02-11T17:45:45 to=2016-02-11T18:15:45 sourcetype=datastore.ps_signaling_records | pivot-table avg(duration) by bearer.apn and bearer.rat-type



	.	EUTRAN (6)	GERAN [2]	UTRAN [1]
	AVG(GTP.DURATION)	AVG(GTP.DURATION)	AVG(GTP.DURATION)	AVG(GTP.DURATION)
.	1.39M	2.29M	4.49M	2.85M
[redacted]				5.15K
[redacted]			11.52M	22.21M
[redacted]		602.78K		6.33
[redacted]				4.32K
[redacted]			18.39M	22.57M
[redacted]		1.14M		3.21M
[redacted]			1.45M	560.86K
[redacted]		1.04M	3.19M	409.29K
[redacted]		998.88K	1.95M	627.62K
[redacted]		1.01M		
[redacted]		523.91K		148.16K
[redacted]		496.34K		
[redacted]		1.07K		
[redacted]		3.87K		
[redacted]				59.66M
[redacted]			1.11M	
[redacted]			1.26M	1.1M
[redacted]			1.12M	
[redacted]			2.27M	
[redacted]				48.72M
[redacted]	944.29K		23.17K	51.38K
[redacted]		86.94K	442.13K	298.63K
[redacted]			5.37M	54.31M
[redacted]	1.06M		12.54M	21.29M

3.6.17 unique

The unique function will return the number of unique keys for the given time range after all filters have been applied.

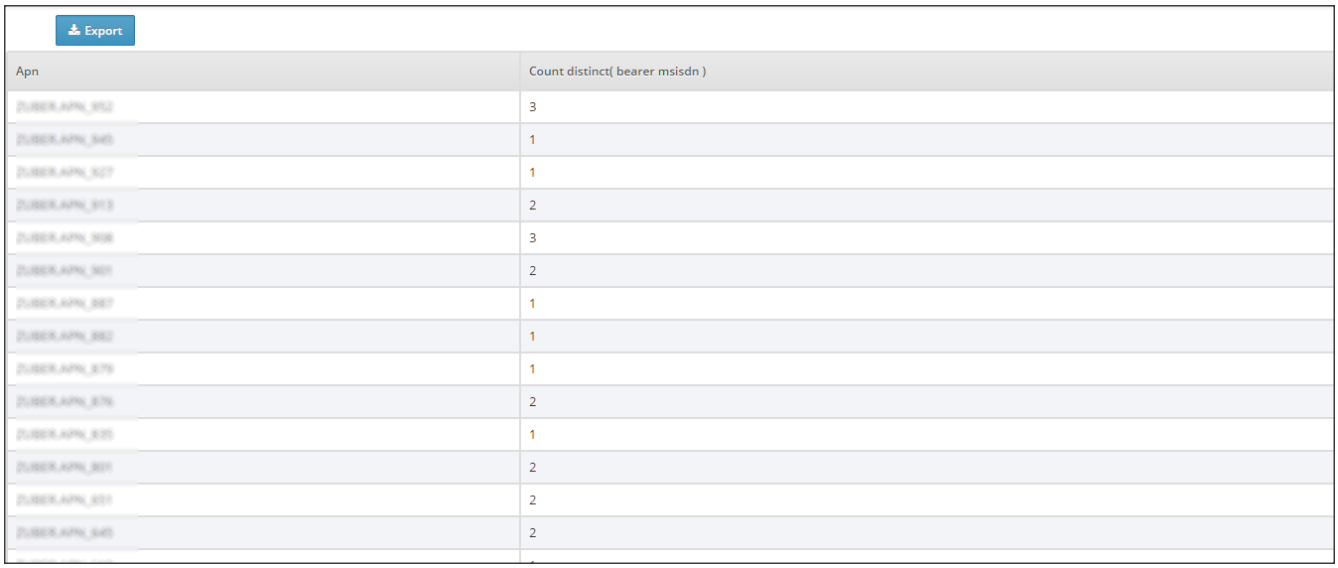
Syntax

unique [field]

*field i*s one of the table field specified in sourcetype. This field is used for aggregation.

Example: Number unique users per APN.

from=2016-02-11T17:45:45 to=2016-02-11T18:15:45 sourcetype=datastore.ps_signaling_records | unique apn



3.6.18 where

The where command allows to specify the conditions to perform the search.

Syntax

where condition

Parameters

condition can be one of the following elements:

1. fieldA operator conditionA

    - fieldA is one of the table field specified in sourcetype.
    - operator can be any of the following operators:
        - < (less than)
        - (greater than)
        - <= (less than or equal)
        - >= (greater than or equal)
        - = (equal)
        - != (not equal)
        - starts-with (starts with for strings)
        - !starts-with (not starts with for strings)
        - ends-with (ends with for strings)
        - !ends-with (not ends for strings)
        - contains (contains for strings)
        - !contains (not contains for strings)
    - conditionA can be:
        - an alphanumeric value. For *example*:

... | *where rat-type = 6*

        - a decorated or translated value. For *example*:

... | *where rat-type = UTRAN*

- a substring of the value in quotes (can contain blank spaces). For *example*:

... | *where device = "Apple iPhone 5"*

... | *where imsi != ""*

2. fieldA in conditionA conditionN

- fieldA can be:
  - one of the table field specified in sourcetype.
  - one of the following reserved words.
  - carrier. to filter by carrier.
  - country. to filter by country.
  - device. to filter by device type.
  - uli-cell. to filter by user location.
-
- conditionA..conditionN are alphanumeric values. For *example*:

... | *where rat-type in 6 4*

3.6.19 reduce

Sometimes it is not enough to perform a simple operation to get some reports but further complexity needs to be added to the query.

The Reduce function provides a great flexibility for the combination of different indicators and expressions to get complex KPI reports.

Syntax

|reduce \[$expression \] [as $expressionName]? [by [$dimension]+]? [-summarization=$sumFunc -time-window=$TIME_WINDOW]?

Parameters

$expression: This is the Reducer expression.

- Expression have the following syntax:
- [$aggregationFunction($expression)]? [$operator]? [$columnOperation]?
-
- Where:
  - *$aggregationFunction*: It is an aggregation function:
- SUM,AVG,MAX,MIN.
  - *$expression*: Expression created by some prior function. E.g: chart
  - *$operator*: Some operator such us '+', '-'. '*', '/'. Note: you can also use mole operators such as 'safediv, min($element1, $element2), etc.
  - *$columnOperation*: It is an operation performed over columns instead of rows. There are several ones (explained later):

- UNIQUE
- BUCKET_UNIQUE
- PERCENTAGE
- PERCENTILE

$expressionName: Expression name to be used in widgets and result tables

$dimension: Dimension(s) used to compose new response.

$TIME_WINDOW: time-window in the same format as timeslice parameter.

$sumFunc: [SUM | AVG | MAX | MAX_BUCKET | MIN | MIN_BUCKET]

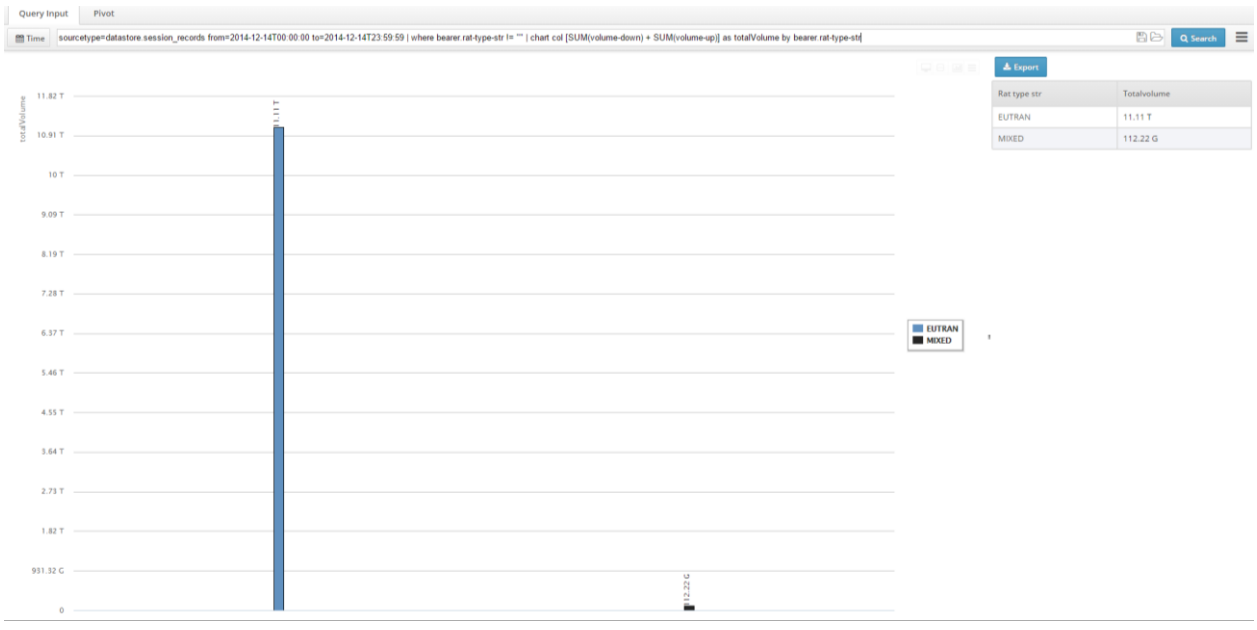
Reduce concept

This section describes how to use reduce operations in MOLE. Reduce operations are in charge of perform both merge and summarization operations over records already aggregated (i.e records returned by some query)

From now on, query response will be named as RESPONSE and reduce-function response as REDUCTION.

Example of RESPONSE (when we do not use reduce clause):

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:59 | where bearer.rat-type-str != "" | chart col [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.rat-type-str



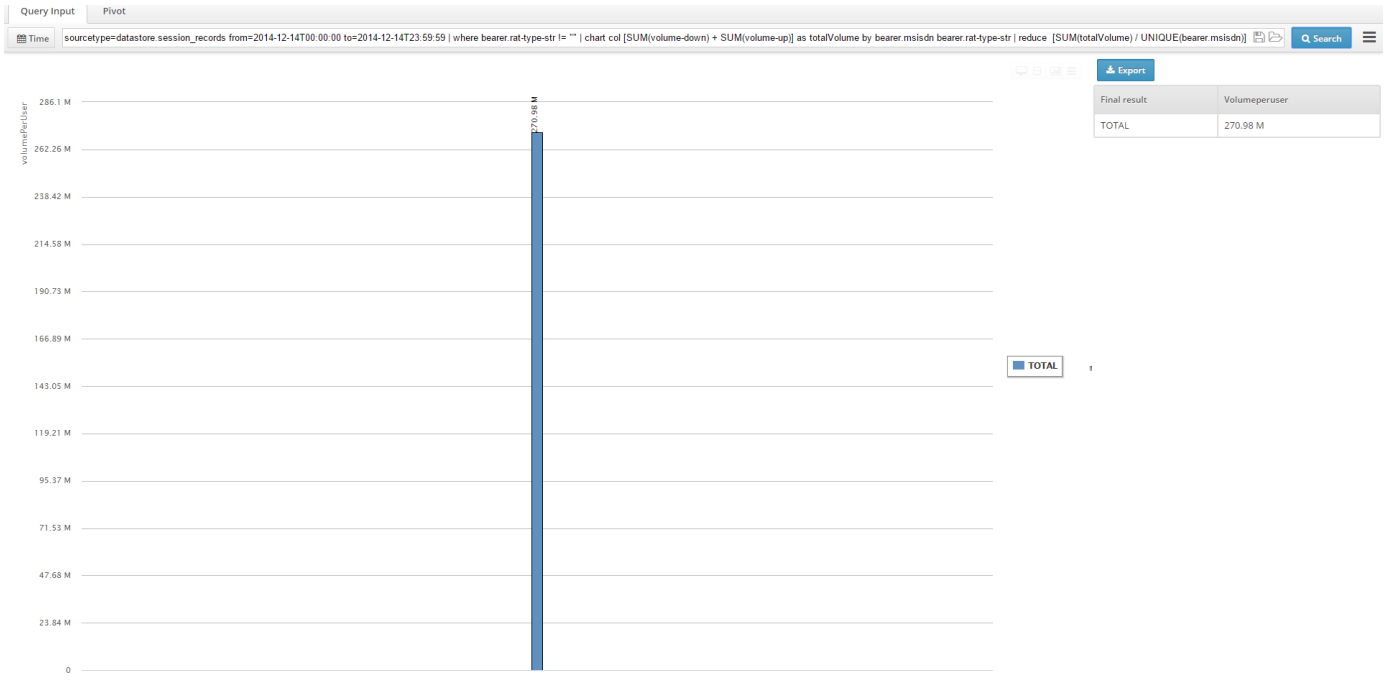
Example of REDUCTION:

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:59 | where bearer.rat-type-str != "" | chart col [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn bearer.rat-type-str | reduce [SUM(totalVolume) / UNIQUE(bearer.msisdn)] as volumePerUser by bearer.rat-type-str

There are 3 different kind of reductions:

- One Dimension and One Value: The REDUCTION will be one row with format “TOTAL-VALUE” and it will be the result of combining the different operations defined in the reduce-function. In order to get this reduction, “by” clause must not be defined. For *example*:

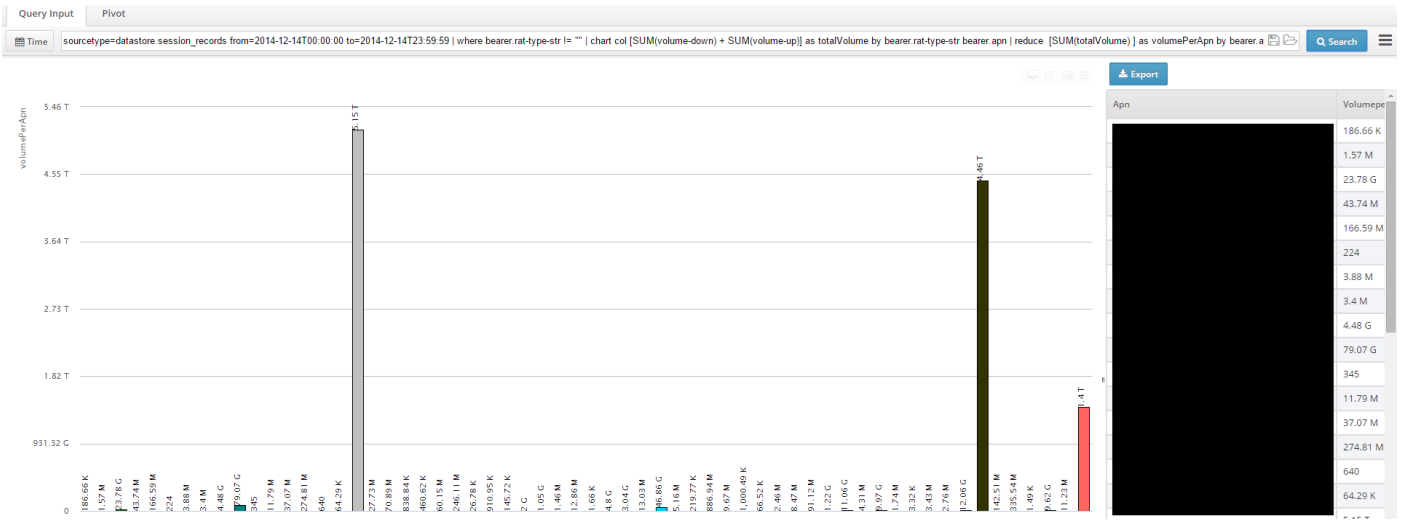
sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:59 | where bearer.rat-type-str != "" | chart col [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn bearer.rat-type-str | reduce [SUM(totalVolume) / UNIQUE(bearer.msisdn)] as volumePerUser



- Combining dimensions: The REDUCTION will be one row per each defined new dimension in REDUCTION. New dimensions are created if "by" clause has less dimensions than *chart* function and then new metrics are created for each of them. For *example*:

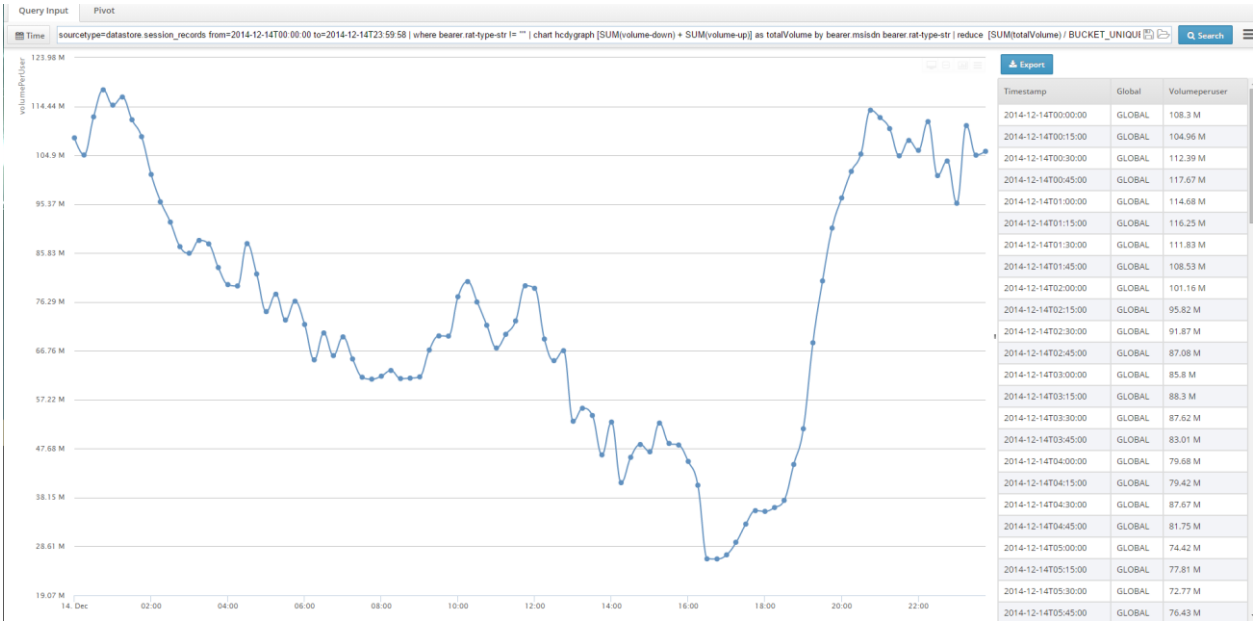
sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:59 | where bearer.rat-type-str != "" | chart col [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.rat-type-str bearer.apn | reduce [SUM(totalVolume)] as volumePerApn by ***bearer.apn***

In this case, note that bearer.rat-type-str dimension has been removed and thus each column represents one apn.



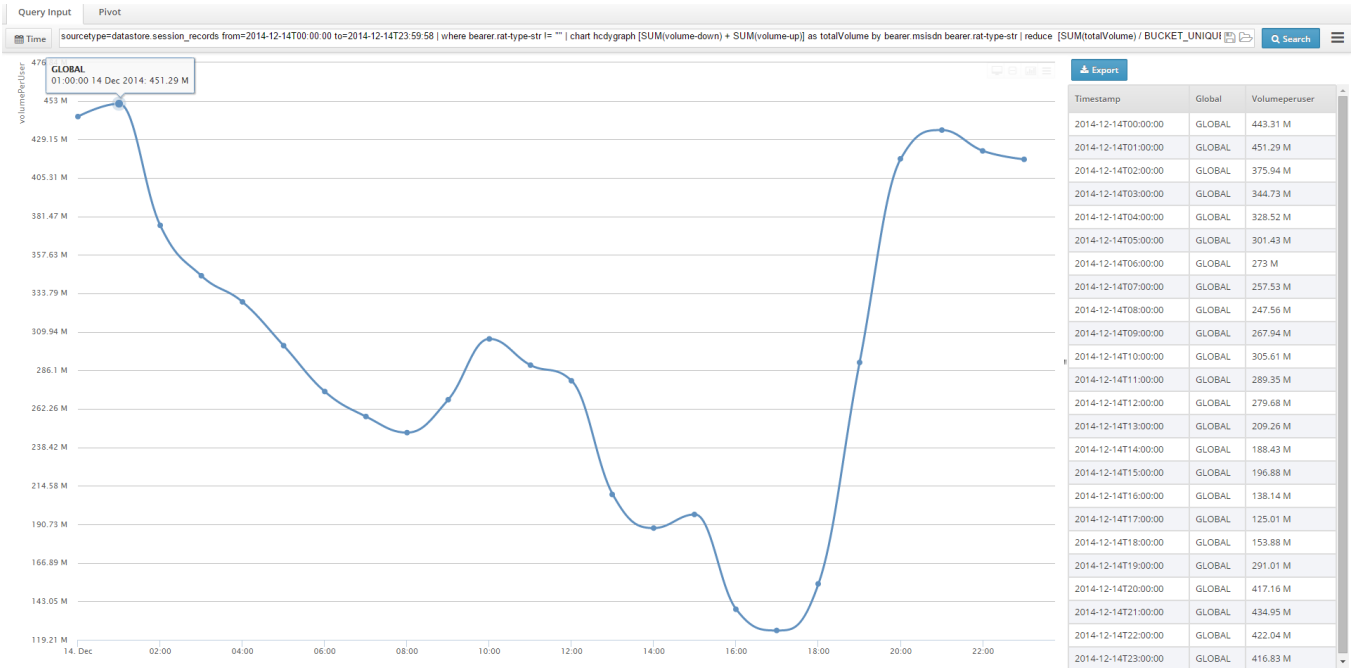
- Timelines: The REDUCTION will be presented as a timeline, that is, each record will contain some period representing some bucket. For *example*:

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | where bearer.rat-type-str != "" | chart hcdygraph [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn bearer.rat-type-str | reduce [SUM(totalVolume) / ***BUCKET_UNIQUE****(bearer.msisdn)] as volumePerUser by* ***timestamp***



Timelines can be summarized using the parameter –time-window (explained later). This is just an *example*:

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | where bearer.rat-type-str != "" | chart hcdygraph [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn bearer.rat-type-str | reduce [SUM(totalVolume) / BUCKET_UNIQUE(bearer.msisdn)] as volumePerUser by timestamp -time-window=01:00:00 -summarization=SUM



Reduce stages

Reduce function can be explained in several stages:

1º Calculation of "Filter operations":

- Filter operations are those ones that do not create new metrics and dimensions but only filter records by using some filter type. Nowadays, there are two filter operations:
  - TOP($indicator,$numberOfElements, $summarizationFunction, $otherCategory, $dimensionList):
    - It calculates tops for the defined $indicator.
    - Number of tops is defined by $numberOfElements.
    - $summarizationFunction is used to aggregate values for different records (useful for timelines as you want to know the aggregated value).
    - $otherCategory is used to group non-top elements in one record.
    - $dimensionList: It is useful to define over which dimensions you want to extract tops.
  - BOTTOM($indicator,$numberOfElements,$summarizationFunction, $otherCategory, $dimensionList):
    - Calculates bottoms for the defined $indicator.
    - Number of bottoms is defined by $numberOfElements.
    - $summarizationFunction is used to aggregate values for different records (useful for timelines).
    - $otherCategory is used to group non- bottoms elements in one record.
    - $dimensionList: It is useful to define over which dimensions you want to extract bottoms.

2º Calculation of "Column operations":

- Column operations are those ones that create some value based on the values for some column of RESPONSE. Nowadays, there are three:
  - UNIQUE(column1, $dimensions) → Compute the unique values for column1 over the defined $dimensions. If no dimensions are defined, there will be only a GLOBAL dimension. This operation adds an extra column containing the unique result.
  - BUCKET_UNIQUE (column1, $dimensions) → Compute the unique values for column1 over the defined $dimensions in each bucket. If no dimensions are defined, there will be only an only dimension named GLOBAL. It needs to be a timeline RESPONSE. This operation adds an extra column containing the unique result.
  - PERCENTAGE(column1) → Compute the percentage of each column value over the total column value (that is, the sum of all values for column1). This operation replaces column1 by the new value.
  - BUCKET_PERCENTAGE (column1) → (note: same as prior one) Compute the percentage for each column value for each bucket. It needs to be a timeline RESPONSE. This operation replaces column1 by the new value.
  - PERCENTILE(column1, $percentileNumber, $dimensions) : Computes the percentile specified by $percentileNumber (e.g 95) for the defined dimensions. This operation returns only the record representing the defined percentile.
  - BUCKET_PERCENTILE(column1, $percentileNumber, $dimensions) : Computes the percentile specified by $percentileNumber (e.g 95) for the defined dimensions in each bucket. It needs to be a timeline RESPONSE. This operation returns only the record representing the defined percentile.

3º Aggregation of each function within the REDUCTION expression (e.g SUM(eventCount)).

- Each RESPONSE formula will be aggregated (e.g SELECT [COUNT(message-type)] as eventCount) by the defined aggregation function (e.g REDUCE [SUM(eventCount)]).
- Available aggregations:
  - SUM -> Sum all values for each dimension
  - AVG -> Average of all values for each dimension
  - MAX -> Max of all values for each dimension
  - MIN -> Min of all values for each dimension

4º Formula evaluation

- Formula defined in reduce-function will be applied for each aggregated record and column operations (if any).

5º Summarization of response created in step 3 (Only if –time-window is defined).

- When a time-window is defined in a timeline RESPONSE, REDUCTION is summarized using –time-window param. It is important to notice that RESPONSE timeslice must be less than REDUCTION time-window.
- Available aggregations:
  - SUM -> Sum all values for each dimension
  - AVG -> Average of all values for each dimension
  - MAX -> Max of all values for each dimension
  - MIN -> Min of all values for each dimension
  - MAX_BUCKET -> Same as MAX but it keeps as dimension the bucket where the max has been found for each dimension. An *example* will be shown later.
  - MIN_BUCKET -> Same as MIN but it keeps as dimension the bucket where the min has been found for each dimension. An *example* will be shown later.

Reduce dimension semantic

Special attention must be paid in how to combine dimensions. That is, there are no lots of restrictions in how to define *"by"* clause and $dimension parameter for column/filter operations. That is, you can combine dimensions as whatever you want (*) just knowing what is the meaning of such combination.

In order to explain this, dimension combinations are done in a simple result table:

ras-client	apn	Total Volume
GGSN1	PREPAID	4 TB
GGSN1	POSTPAID	2 TB
GGSN2	PREPAID	5 TB
GGSN3	PREPAID	3 TB
GGSN4	POSTPAID	4 TB

Prior table could be obtained by this chart function:

| chart summary-table [SUM(volume)] as totalVolume by ras-client apn

Now, if you want to calculate the top 1 Ras-Client, you can use reduce like this:

| reduce [TOP(totalVolume, 1, SUM, , ras-client apn)] as topOne by ras-client apn

In this operation, reduce will sum all total volume's for each RAS-CLIENT and APN and will return this result:

ras-client	apn	Total Volume
GGSN2	PREPAID	5 TB

As you can see, we have defined dimensions ras-client and apn twice in order to get a meaningful result:

- Once in TOP operation
- Once in "by" clause.

Normally, dimensions for column/filter operations and reduce dimensions will be the same, just note that sometimes you can "play" with this in order to get whatever you want.

In later sections you will fine each operation sintax.

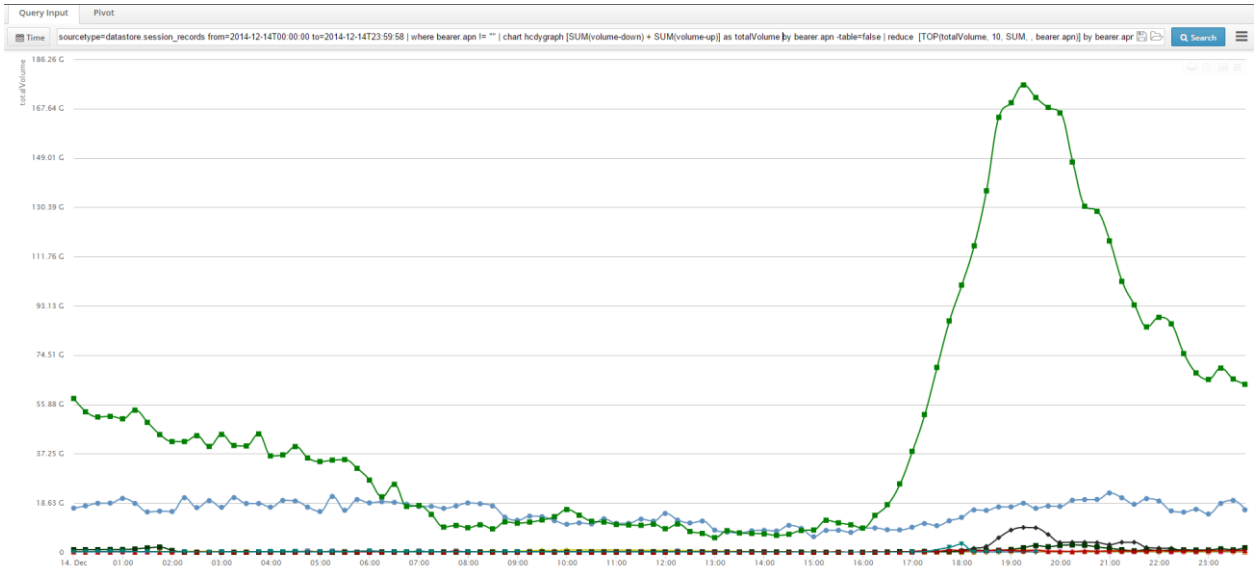
(*) There is an important restriction for filter operations: "by" clause in *reduce* must preserve the same dimensions as the response coming from another function (e.g chart). This restriction is caused by the fact that filter operations neither modify dimensions nor metrics.

Reduce filter operations

TOP Function *Examples*:

- *Example*1: Calculate top-10 APNs by volume. As this indicator is a sum, use SUM as summarization function.

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | where bearer.apn != "" | chart hcdygraph [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.apn -table=false | reduce [TOP(totalVolume, 10, **SUM**, , bearer.apn)] by bearer.apn

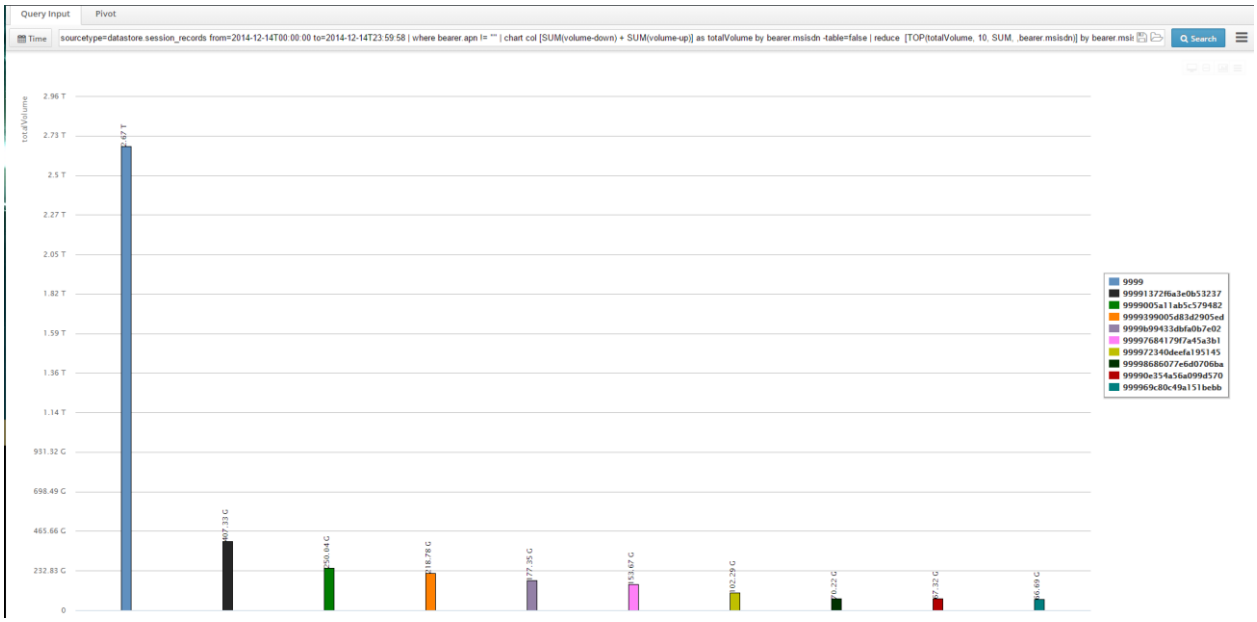


- *Example*2: Calculate top-10 APNs and group non-top ones in a record named OTHER

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | where bearer.apn != "" | chart hcdygraph [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.apn -table=false | reduce [TOP(totalVolume, 10, SUM, ***OTHER*** *, bearer.apn)] by bearer.apn*

- *Example*3: Calculate top-10 individual contributors by volume. As this indicator is a sum, use SUM as summarization function.

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | chart col [SUM(volume-down) + SUM(volume-up)] as totalVolume by ***bearer.msisdn*** *-table=false | reduce [TOP(totalVolume, 10, SUM, ,bearer.msisdn)] by* ***bearer.msisdn***



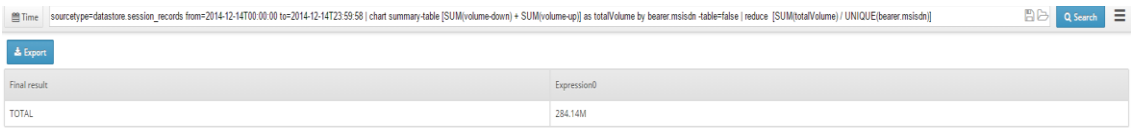
BOTTOM *Function Examples*: It can be used prior queries by replacing TOP by BOTTOM.

Note: For filter operations, *"chart by"* and *"reduce by"* clauses must have the same dimensions. For *example*:

chart hcdygraph [SUM(volume-down) + SUM(volume-up)] as totalVolume **by bearer.apn** -table=false | reduce [TOP(totalVolume, 10, SUM, **OTHER** , bearer.apn)] **by bearer.apn**

Reduce column operations

- UNIQUE:
- Get the volume for each user and then sum it and divide it by unique users in that response:
  - Get the volume for each user:
    - chart summary-table [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn
  - Then sum it:
    - chart summary-table [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn | reduce [SUM(totalVolume)]
  - And divide it by unique users:
    - chart summary-table [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.msisdn | reduce [SUM(totalVolume) / UNIQUE(bearer.msisdn)]
- Result:

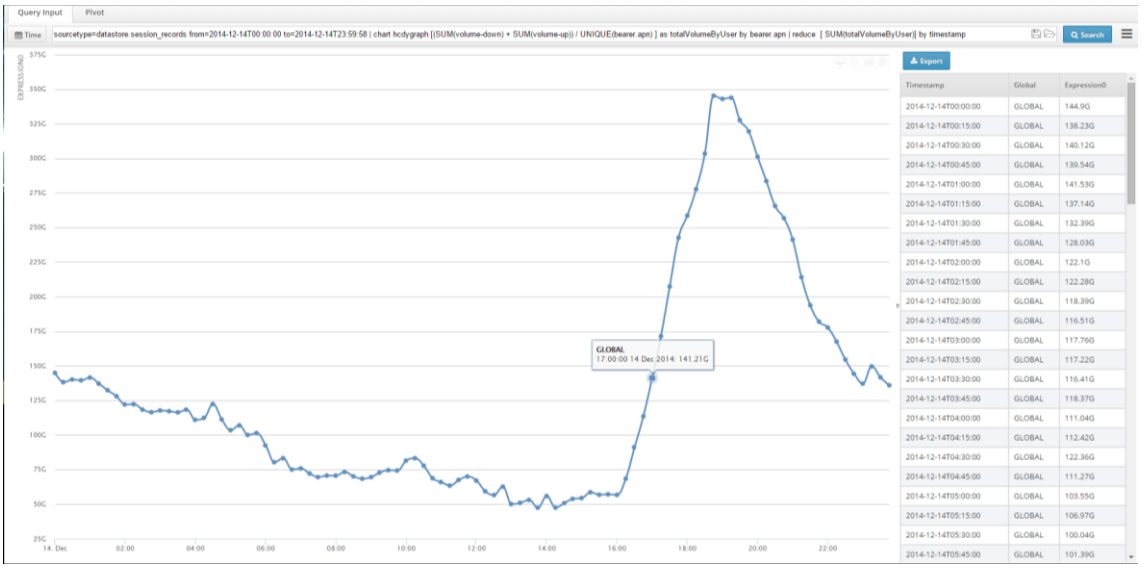


- BUCKET_UNIQUE:

Get the volume divided by unique users per APN each 15 minutes and sum that result each 15 minutes.

- Get the volume divided by unique users per APN:
  - chart hcdygraph [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn
- And sum that result each 15 minutes
  - chart hcdygraph [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn | reduce [SUM(totalVolumeByUser)] **by timestamp**

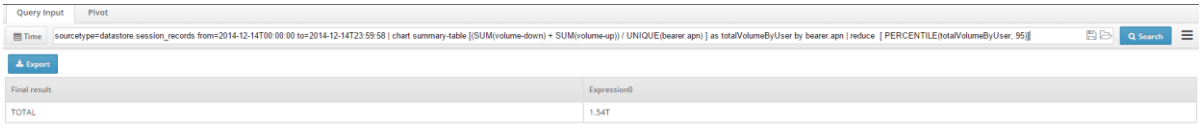
Result:



- PERCENTILE:

Get the volume divided by unique users for each APN and return the value which represents the percentile 95:

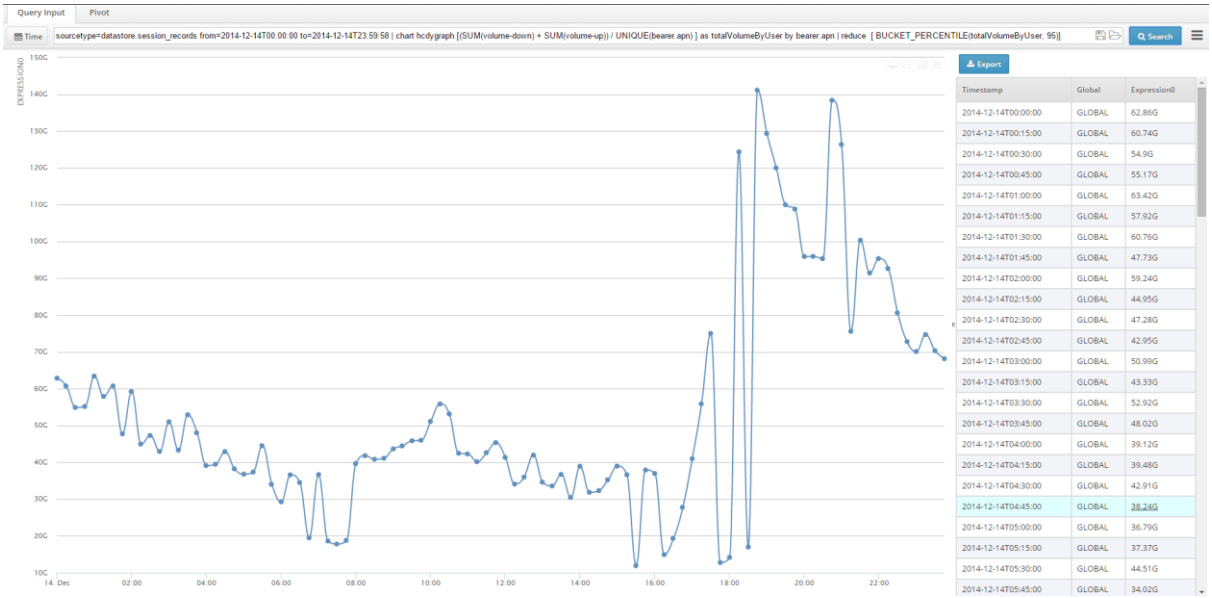
sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | chart summary-table [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn | reduce [***PERCENTILE****(totalVolumeByUser, 95)]*



- BUCKET_PERCENTILE:

For each 15 minutes, get the volume divided by unique users for each APN and return the value which represents the percentile 95:

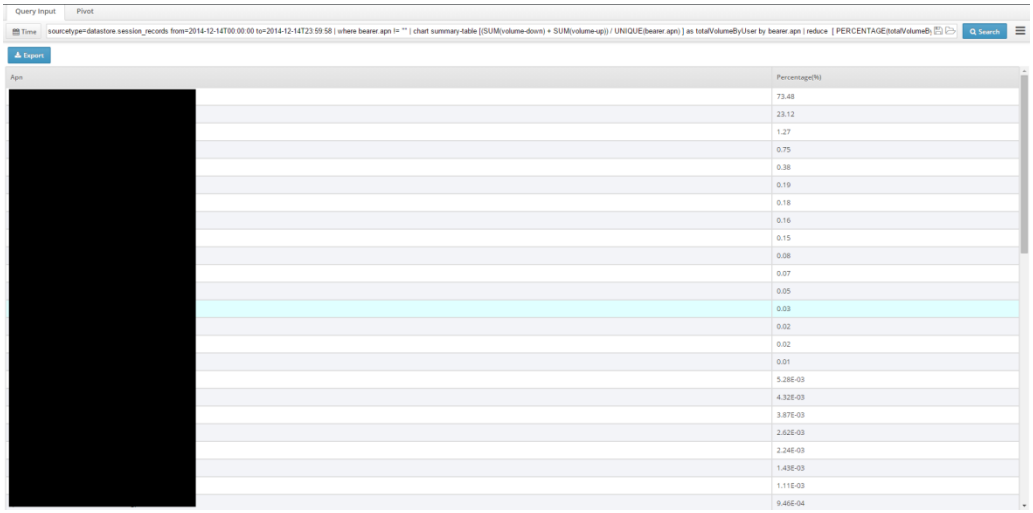
sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | chart dygraph [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn | reduce [***BUCKET_PERCENTILE****(totalVolumeByUser, 95)]*



- PERCENTAGE:

Get the total volume divided by unique users and modify response in order to get which percentage represents each apn.

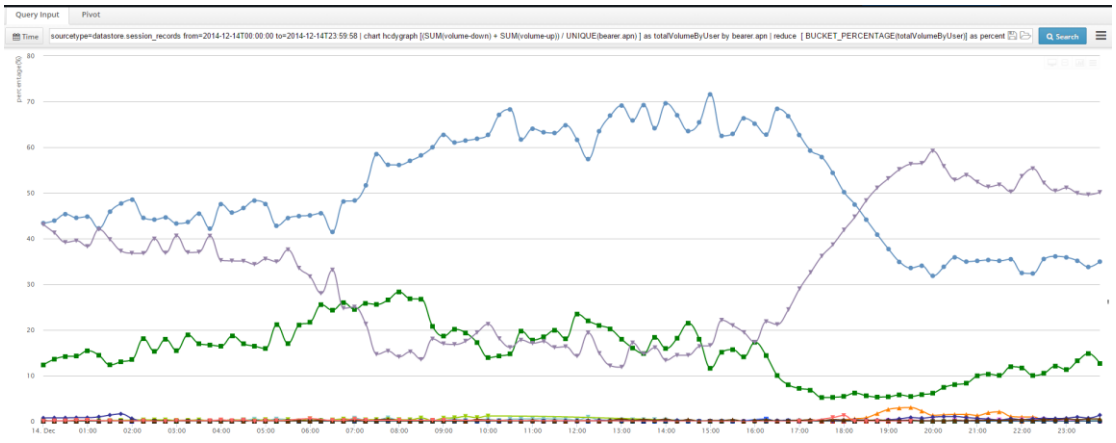
sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | chart summary-table [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn | reduce [***PERCENTAGE****(totalVolumeByUser)] as percentage(%) by bearer.apn*



- BUCKET_PERCENTAGE:

Each 15 minutes, get the total volume divided by unique users and modify response in order to get which percentage represents each apn.

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | chart hcdygraph [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn | reduce [***BUCKET_PERCENTAGE****(totalVolumeByUser)] as percentage(%) by timestamp bearer.apn*



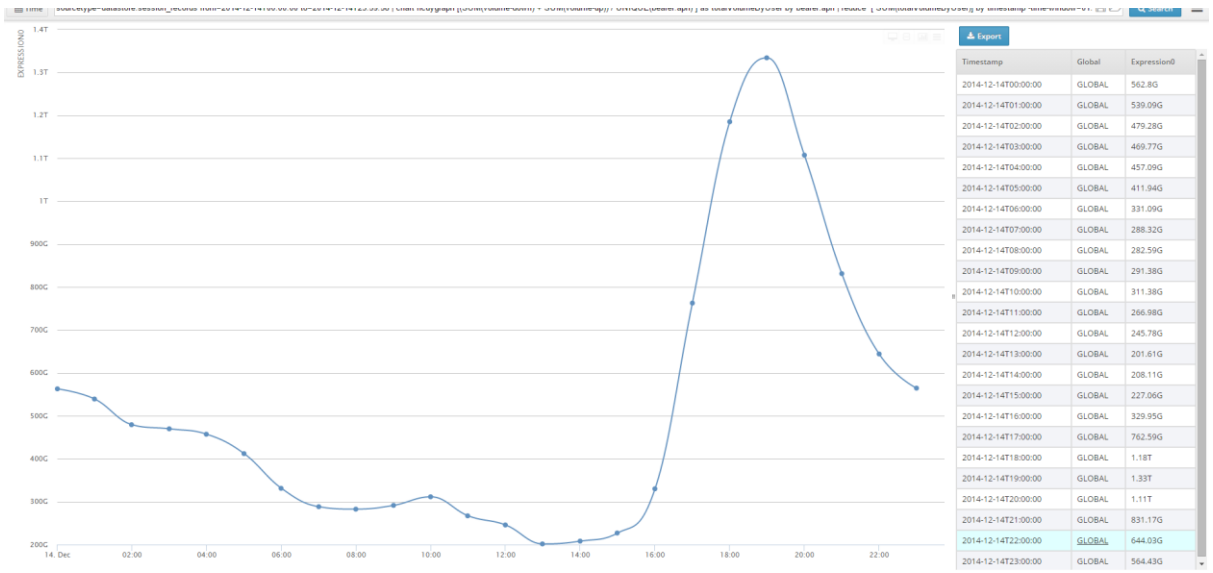
Reduce examples

This section shows a couple of reduce examples with a short explanation.

Example 1 (time-window and summarization options):

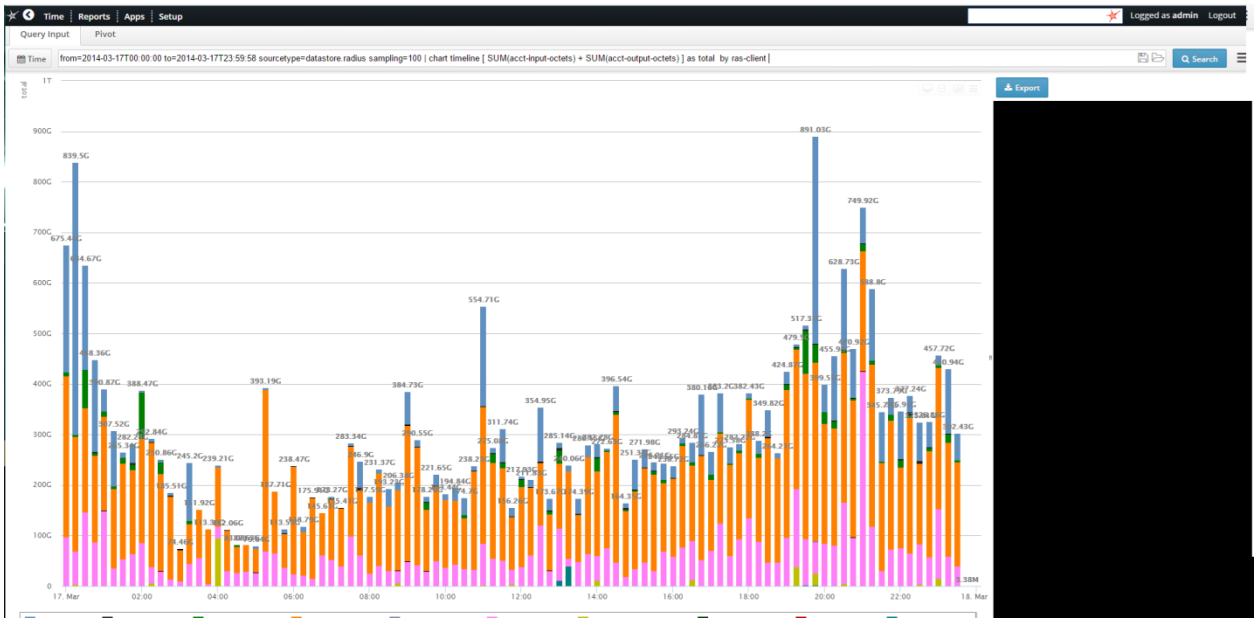
Get the volume divided by unique users per APN each 15 minutes and sum that result summarized by one hour, use SUM as summarization function for each 15-minute value.

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | chart hcdygraph [(SUM(volume-down) + SUM(volume-up)) / UNIQUE(bearer.apn)] as totalVolumeByUser by bearer.apn | reduce [SUM(totalVolumeByUser)] by timestamp -time-window=01:00:00 -summarization=SUM



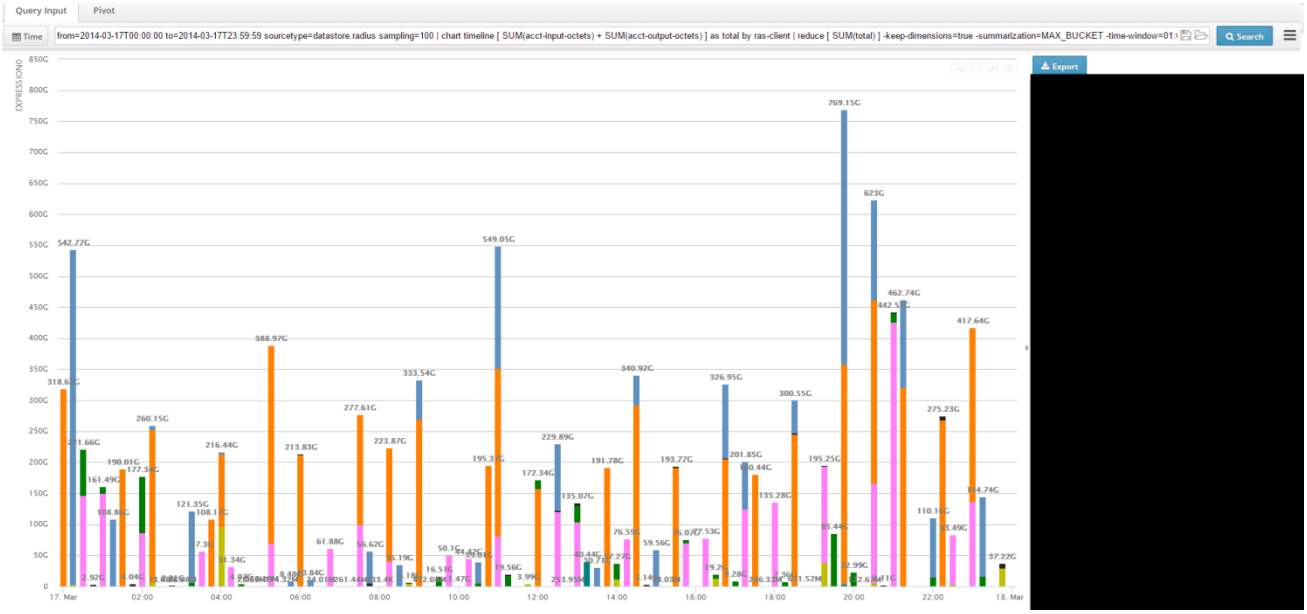
Example 2 (max_bucket and min_bucket summarizations):

Having this report:



Get the 15-minute bucket in a one-hour time-window in which the max total volume is found **per each ras-client**.

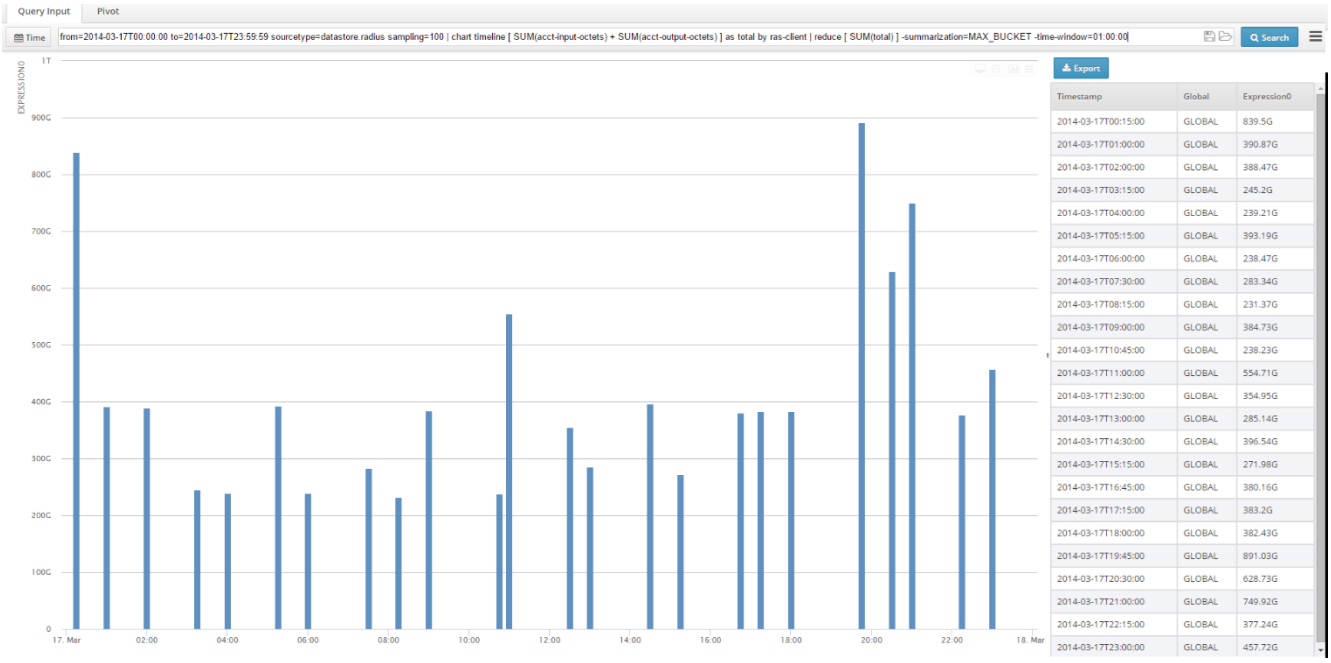
from=2014-03-17T00:00:00 to=2014-03-17T23:59:59 sourcetype=datastore.radius sampling=100 | chart timeline [SUM(acct-input-octets) + SUM(acct-output-octets)] as total ***by ras-client*** *| reduce [SUM(total)] by ras-client -summarization=MAX_BUCKET -time-window=01:00:00*



Instead of getting the max value for ras-client, first aggregate each ras-client volume using sum and then find the 15-minute bucket in a one-hour time-window in which the total max value is found:

from=2014-03-17T00:00:00 to=2014-03-17T23:59:59 sourcetype=datastore.radius sampling=100 | chart timeline [SUM(acct-input-octets) + SUM(acct-output-octets)] as total by ras-client | reduce [SUM(totalVolume)] -summarization=MAX_BUCKET -time-window=01:00:00

(Note that by clause has been removed)



Same applies for MIN_BUCKET. Just replace MAX_BUCKET by MIN_BUCKET.

Example 3 (Nested reduce's):

It is possible to defined several reduces in order to modify several times the main response.

For *example*: Get the top 10 APN's by volume and then, represent those ones in percentages:

sourcetype=datastore.session_records from=2014-12-14T00:00:00 to=2014-12-14T23:59:58 | where bearer.apn != "" | chart summary-table [SUM(volume-down) + SUM(volume-up)] as totalVolume by bearer.apn | reduce [TOP(totalVolume, 10, SUM, , bearer.apn)] by bearer.apn | reduce [PERCENTAGE(totalVolume)] as topVolume(%) by bearer.apn

3.7 Useful tips

This section will show some useful tips in order to perform advance full search operations.

3.7.1 Reduce

- If you want to use summary-table in reduce functions be aware that it only will apply over the first 100 registers unless you use the clause "| head 0".

  Example:

  title="Reduce by MAX" sourcetype=datastore.nodes_records from=2014-10-01T00:00:00 to=2014-10-25T23:59:59 timeslice=01:00:00 | chart summary-table [SUM(volume)/ SUM(cpu)] as relation by node | **head 0** | reduce [MAX(relation)]

- Use the percentile and mavg functions with 2 parameters (not 3 because it doesn't work such as we expect).

  Example:

  sourcetype=datastore.pdu_streaming_records timeslice=01:00:00 from=2016-06-09T12:00:00 to=2016-06-09T12:59:59 | where event.text = PDU | where event.interim-id = 0 | where net.peak-throughputdown > 1000 | where stream.duration > 0 | where stream.effective-reproduction-time > 4 | where http.setup-time >= 0 | chart table [AVG(tcp.service-start-time)] as temp1 by imsi | where bearer.imsi starts-with 2200110098 | **reduce [PERCENTILE(temp1, 80)]** as temp2 by timestamp