Monitor the usage of volumes and forecast the requirements for their integrated cloud data services

Case Study, IQLect

AT A GLANCE

1. Overview
2. The Problem
3. The Solution
4. Conclusion

OVERVIEW

NetApp is a cloud data services and data management company. NetApp offers hybrid cloud data services for management of applications and data across and on-premises environments.

They Provide data space or volumes to their users. They have space containers and every container has some fixed space. Each container has many clusters. Their users are grouped in these clusters, every cluster has some fixed memory allotted to it. They wanted a solution that can monitor the data usage pattern for users and can predict when the space allotted will be full so that they can increase storage before it goes down due to memory shortage.

THE PROBLEM

To monitor data usage of their cloud services (volumes) and to send alerts if there is any issue detected . Also forecast the usage for coming days for proactive and timely provisioning.

THE SOLUTION

Creating Schema

The raw data contains cluster id, the Hapair and aggregate for each user as well as their total data capacity, used data capacity, their percentages, available data capacity, daily growth rate percentage and few other attributes.

The raw stream schema is given below.

```
{
   "name": "data_stream",
   "type": 1, "inpt": [],
   "attr": [ { "name": "cluster","type": 5,"kysz": 64,"sidx": 1,"stat":2"},
    { "name": "hapair","type": 5,"kysz": 64,"sidx": 1,"stat": 2 },
    { "name": "aggregate","type": 5,"kysz": 64,"sidx": 1, "stat": 2},
    { "name": "totaldatacapacity","type": 11 },
    { "name": "useddatacapacity","type": 11, "stat": 3},
    { "name": "useddatapercent","type": 11 },
    { "name": "availabledatacapacity","type": 11 },
    { "name": "availabledatapercent","type": 11 },
    { "name": "growthrate","type": 11 },
    { "name": "type","type": 5}, { "name": "raid_type","type": 5 },
    { "name": "aggregate_state","type": 5},
    { "name": "snaplock_type","type": 5},
    { "name": "date","type": 5}
    ],
   "swsz": 86400
   }
 ]
}
```

Once schema is registered, we can start building our monitoring app

To view schema:

1. Creating unique id for each user

Structure to identify particular user is that we have to look for cluster then their hapari id and then aggregate, here we are creating a unique id for every user.

Adding catr in data_stream :

```
"catr" : [
  {
  "fnr" : 1,"type" : 5, “seq”: 1,”stat”: 2,
  "name" : "vid",
  "opnm" : "ADD",
  "iatr" : [ “cluster”,“hapair","aggregate" ] }
]
```

This will generate unique value for each user

Query to check total unique users:

```
bangdb> select aggr(vid) from netapp.data_stream
```

2. Monitor the average data used capacity and growth rate for every cluster-hapari-aggregate

To achieve this : Add group-by in the mainstream where we are computing average data used capacity based on Cluster-Hapair-Aggregator.

Adding Two groupby one for used data capacity and other for daily growth rate :

```
"gpby": [
  { "iatr": "useddatacapacity",
  "gpat": ["cluster",”hapair”,”aggregate” ],
  "kysz": 124, "gran": 3600, "stat": 3
  },
  { "iatr": "growthrate",
  "gpat": ["cluster",”hapair”,”aggregate” ],
  "kysz": 120, "gran": 3600, "stat": 3
  }
]
```

Query to get results:

```
bangdb> select aggr(useddatacapacity) from netapp.data_stream groupby
cluster:hapair:aggregate
```

```
bangdb> select aggr(growthrate) from netapp.data_stream groupby
cluster:hapair:aggregate
```

3. List all aggregates with used data capacity greater than 85 percent for its total data capacity and monitor their data usage

To achieve this: there are two ways one is that we write a query to get the list on the screen or we create a separate stream to collect details of all aggregate crossing the 85% threshold.

Step 1: calculate the used_data_capacity_percentage

Adding catr to data_stream :

```
"catr" : [
  {
  "fnr" : 1,"type" : 11
  "name" : "useddatapercent",
  "opnm" : "PERCENT",
  "iatr" : [ “useddatacapacity","totaldatacapacity" ] }
]
```

Query to get result:

```
bangdb> select * from netapp.data_stream where used_data_percent > 85
```

Step 2: Adding filter to collect details for aggregate with usage above 85 and adding groupby to monitor their average usage.

Adding filter in data_stream with condition usage greater than 85 :

```
“fltr”:[
  {
  "name": "aggr_above_85",
  "fqry": {
  "name": "{\"query\":
[{\"key\":\"useddatapercent\",\"cmp_op\":0,\"val\":85}]
  ,\"qtype\":1}","type": 1 },
  "fatr": [ "used_data_capacity", "cluster","hapair","aggregate",
  "growthrate","date","vid"
  ],
  "ostm": "aggr_above_85"
  }
]
```

Adding filter stream to collect data after filter :

```
{
  "name": "aggr_above_85",
  "type": 2, "inpt":["data_stream"],
  "attr": [
  { "name": "cluster","type": 5,"kysz": 64,"sidx": 1,"stat":2 },
  { "name": "hapair","type": 5,"kysz": 64,"sidx": 1,"stat": 2 },
  { "name": "aggregate","type": 5,"kysz": 64,"sidx": 1, "stat": 2},
  { "name": "useddatacapacity","type": 11, "stat": 3},
  { "name": "growthrate","type": 11 },
  { "name": "useddatapercent","type": 11},
  { "name": "vid","type": 5},
  { "name": "date","type": 5}
  ],
  "swsz": 86400
}
```

Step 3: Monitoring average used data capacity

Adding groupby :

```
"gpby": [
  { "iatr": "useddatacapacity",
  "gpat": ["cluster",”hapair”,”aggregate” ],
  "kysz": 124, "gran": 3600, "stat": 3
  },
  { "iatr": "growthrate",
  "gpat": ["cluster",”hapair”,”aggregate” ],
  "kysz": 120, "gran": 3600, "stat": 3
  }
]
```

Query to get result

```
bangdb> select aggr(useddatacapacity) from netapp.aggr_above_85 groupby
cluster:hapair:aggregate
```

```
bangdb> select aggr(growthratepercent) from netapp.aggr_above_85 groupby
cluster:hapair:aggregate
```

4. Send alerts if data usage for an aggregate is above 95 percent and continuously increasing with growth rate above average value.

Here, we have 3 conditions to check before sending a notification. First the usage should be above 95, second usage should be above 95 more then once (should continuously happen) and third growth rate should be high. For this we will use cep, it's perfect for cases like this one.

Adding cep of type 1 in above filter stream :

```
"cepq": [
  {
  "name": "agger_above_95","type": 1,
  "iatr": ["useddatacapacity","useddatapercent"],
  "rstm": "data_stream",
  "ratr": ["cluster","hapair","aggregate" ],
  "jqry": {
  "cond": ["vid","useddatapercent"],
  "opid": 14,
  "args": ["vid",”95”],
  "cmp": ["EQ","GT"],
  "seq": 0,
  },
  "ostm": "Critical_aggregates",
  "tloc": 1000,
  "cond": [
  {"name": "NUMT","opid": 1,"val": 2},
  {"name": "DUR","opid": 0,"val": 3600}
  ],
  "notf": 805
  }
]
```

Adding cep output stream :

```
{
  "name": "Critical_aggregates",
  "type": 3, "inpt":["aggr_above_85"],
  "attr": [
  { "name": "cluster","type": 5,"kysz": 64,"sidx": 1,"stat":2 },
  { "name": "hapair","type": 5,"kysz": 64,"sidx": 1,"stat": 2 },
  { "name": "aggregate","type": 5,"kysz": 64,"sidx": 1, "stat": 2},
  { "name": "useddatapercent","type": 11},
  { "name": "useddatacapacity","type": 11, "stat": 3}
  ],
  "swsz": 86400,
  "gpby": [
  { "iatr": "useddatacapacity",
  "gpat": ["cluster","hapair","aggregate" ],
  "kysz": 124, "gran": 3600, "stat": 3
  }
  ]
}
```

Notification id 805 will be generated as soon as this event occurs.

5. List aggregate with slow average growth rate and data usage below 35 percent of total data capacity

We can directly do this using query :

```
bangdb> select * from netapp.data_stream where growth_rate < 2 and
used_data_percent < 35
```

6. Send an alert if there is a sudden increase in data usage for an aggregate.

Adding cep :

```
"cepq": [
  {
  "name": "anomaly",
  "type": 6,
  "tloc": 1000,
  "fqry": {
  "type": 1,
  "name":
  "{\"query\":
[{\"key\":\"useddatacapacity\",\"cmp_op\":0,\"val\":\"avg(data_stream.used_d
ata_capacity,h_1,more_50)\"}],\"qtype\":3}"
  },
  "notf": 806
  }
]
```

7. List aggregate with negative growth rate.

Query to achieve this :

```
bangdb> select * from netapp.data_stream where growth_rate < 0
```

8.Training model to predict next day data usage.

We will be predicting the usage for aggregates with usage capacity already above 25% of total data capacity. Dataset contains all aggregates with usage data above 25% as in the dataset given to us, we don't have any aggregate above 30%.

Steps to train time series model

1. Enter the command “train model netapp_1”

   [on entering above command cli workflow starts for ml training]

   Following is the workflow for ml training, the user just have to enter the training details
2. what's the name of the schema for which you wish to train the model?: netapp

   [enter the schema/app name]
3. do you wish to read earlier saved ml schema for editing/adding? [yes | no]: no

   From the list of algorithms for time series, user have to select 10
4. what's the algo would you like to use (or Enter for default (1)): 10
5. enter the input file for forecast training (along with full path):

   /home/iql0005/Desktop/SC_test/netapp.csv

   [location of training file on local system]
6. what's the format of the file [CSV (1) | JSON (2)] (or Enter for default (1)): 1

   [format of the training file]
7. what's the position of timestamp field in the csv (count starts with 0): 0

   [provide the position of date/time attribute in the training file]
8. enter the time-stamp field datatype [string (5) | long (9) | double (11)] (or Enter for default (9)): 5

   [here, we have to tell whether time attribute is in which format--if any date format the select 5, if in timestamp then select 11/9]
9. what's the position of entity field in the csv (count starts with 0, enter -1 ignore): 4

   [As we have many aggregates]
10. what's the position of quantity field in the csv (count starts with 0): 6

    [here, used capacity will be target, provide the position of this attribute in training file]
11. do you wish to aggr for higher time dimension? [yes | no]: yes

    [As we have to perform aggregation day wise]
12. what's the lag for data to be selected for prediction (default is 2)?: 6

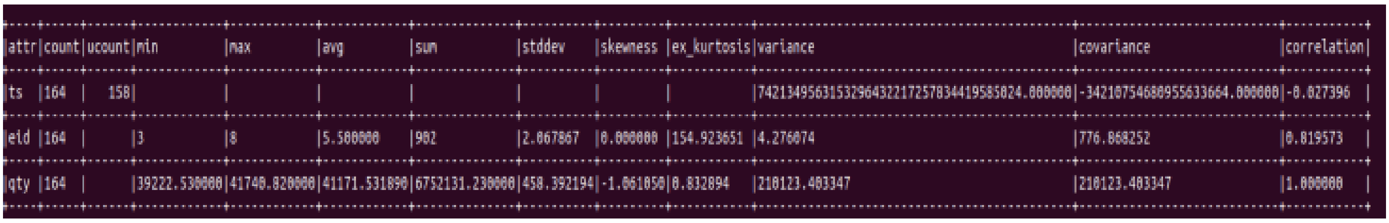
    [to predict next day volume, taking the last 6 days used capacity values. In which present value will be selected as target for training]
13. what's the offset for data to be selected for prediction (default is 0)?: 0

    [As we are predicting next days values]

Once we are done, we get forecast train request :

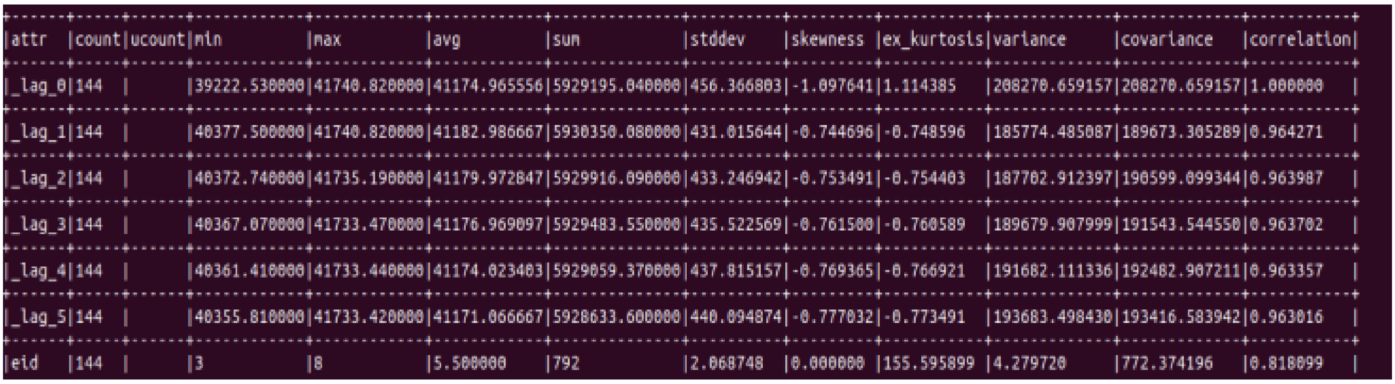
```
{
  "input_file" : "/home/iql0005/Desktop/SC_test/netapp.csv",
  "tsfid" : 0,--------------------position of date/time attribute in training csv file
  "schema-name" : "netapp",-----schema name
  "eidtype" : 9,---------------attribute eid data type
  "model_name" : "netapp_1",----- model name
  "qtydtype" : 11,--------------------data type for quantity/target attribute
  "ignore_aggr" : 1,--------------------------whether want to aggregate or not
  "tsf" : "ts",----------------------time stamp
  "tsdtype" : 5,------------data type for date/time attribute
  "lag" : 6,------------------------number of lags selected
  "qtyfid" : 6,------------------position of target attribute in training csv file
  "offt" : 0,--------------------after how many day to predict
  "ipfmt" : 1,---------------aggregate based on day
  "qtyf" : "qty",--------------name of target attribute
  "eidf" : "eid",------------ name of attribute
  "eidfid" : 4 ----- position of eid attribute in csv file
}
```

Before training, we get a data summary which contains stat and correlation details in the data.



```
+----+-----+------+------------+------------+------------+--------------+----------+---------+-----------+--------------------------------------------+----------------------------+-----------+
|attr|count|ucount|min         |max         |avg         |sum           |stddev    |skewness |ex_kurtosis|variance                                    |covariance                  |correlation|
+----+-----+------+------------+------------+------------+--------------+----------+---------+-----------+--------------------------------------------+----------------------------+-----------+
|ts  |164  |   158|            |            |            |              |          |         |           |7421349563153296432217257834419585024.000000|-34210754680955633664.000000|-0.027396  |
+----+-----+------+------------+------------+------------+--------------+----------+---------+-----------+--------------------------------------------+----------------------------+-----------+
|eid |164  |      |3           |8           |5.500000    |902           |2.067867  |0.000000 |154.923651 |4.276074                                    |776.868252                  |0.819573   |
+----+-----+------+------------+------------+------------+--------------+----------+---------+-----------+--------------------------------------------+----------------------------+-----------+
|qty |164  |      |39222.530000|41740.820000|41171.531890|6752131.230000|458.392194|-1.061050|0.832894   |210123.403347                               |210123.403347               |1.000000   |
+----+-----+------+------------+------------+------------+--------------+----------+---------+-----------+--------------------------------------------+----------------------------+-----------+
```

Next we get summary for extracted time series data:



```
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|attr  |count|ucount|min         |max         |avg         |sum           |stddev    |skewness |ex_kurtosis|variance     |covariance   |correlation|
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_0|144  |      |39222.530000|41740.820000|41174.965556|5929195.040000|456.366803|-1.097641|1.114385   |208270.659157|208270.659157|1.000000   |
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_1|144  |      |40377.500000|41740.820000|41182.986667|5930350.080000|431.015644|-0.744696|-0.748596  |185774.485087|189673.305289|0.964271   |
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_2|144  |      |40372.740000|41735.190000|41179.972847|5929916.090000|433.246942|-0.753491|-0.754403  |187702.912397|190599.099344|0.963987   |
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_3|144  |      |40367.070000|41733.470000|41176.969097|5929483.550000|435.522569|-0.761500|-0.760589  |189679.907999|191543.544550|0.963702   |
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_4|144  |      |40361.410000|41733.440000|41174.023403|5929059.370000|437.815157|-0.769365|-0.766921  |191682.111336|192482.907211|0.963357   |
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_5|144  |      |40355.810000|41733.420000|41171.066667|5928633.600000|440.094874|-0.777032|-0.773491  |193683.498430|193416.583942|0.963016   |
+------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|eid   |144  |      |3           |8           |5.500000    |792           |2.068748  |0.000000 |155.595899 |4.279720     |772.374196   |0.818099   |
```

14. do you wish to schedule for the training of the forecast model now? [yes | no]:

yes

[starting training by saying yes]

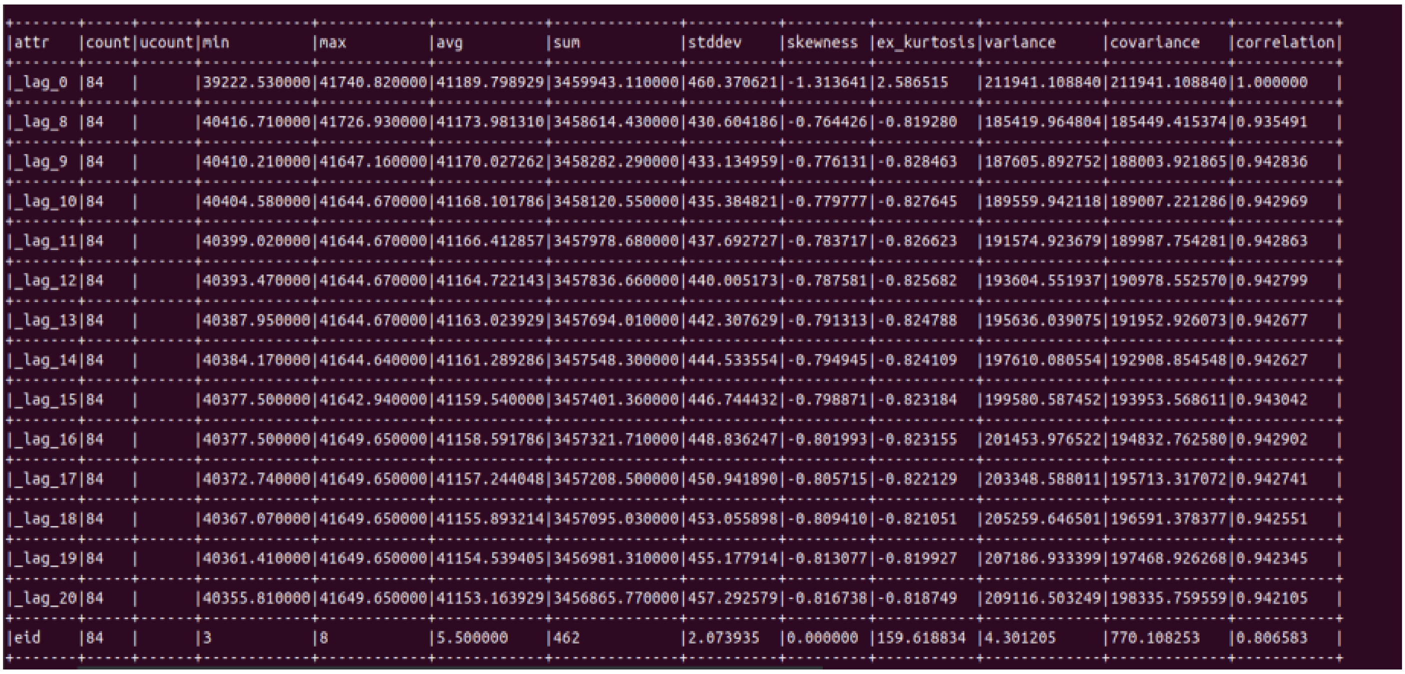
scheduled for the model [netapp] training-- we get this once the training have started on cli

9.Training model to predict usage for the 7th day

Following the same approach as above, we will get

```
{
  "eidtype" : 9,-----------------eid attribute data type
  "input_file" : "/home/iql0005/Desktop/SC_test/netapp.csv",
  "aggr" : 2,-----------------performing aggregation average
  "eidf" : "eid",-----------------name of attribute
  "offt" : 7,-------------------- to predict for 7th day
  "tsdtype" : 5,--------------------time attribute is of date format
  "ipfmt" : 1,
  "qtyf" : "qty",-----------------------name of target attribute
  "qtyfid" : 6,---------------------------position of target attribute in training file
  "lag" : 15,--------------------------------number of lag attributes to be used
  "gran" : 1,-----------------------------aggregating based on day
  "qtydtype" : 11,--------------------------data type of target attribute
  "tsf" : "ts",--------------------------------time stamp
  "model_name" : "netapp7d_15",-----------------model name
  "tsfid" : 0,----------------------------position of date attribute in training file
  "eidfid" : 4, -------------------------position of eid attribute in training file
  "schema-name" : "netapp"-----------------name of schema
}
```

Data summary for second model :



```
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|attr   |count|ucount|min         |max         |avg         |sum           |stddev    |skewness |ex_kurtosis|variance     |covariance   |correlation|
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_0 |84   |      |39222.530000|41740.820000|41189.798929|3459943.110000|460.370621|-1.313641|2.586515   |211941.108840|211941.108840|1.000000   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_8 |84   |      |40416.710000|41726.930000|41173.981310|3458614.430000|430.604186|-0.764426|-0.819280  |185419.964804|185449.415374|0.935491   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_9 |84   |      |40410.210000|41647.160000|41170.027262|3458282.290000|433.134959|-0.776131|-0.828463  |187605.892752|188003.921865|0.942836   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_10|84   |      |40404.580000|41644.670000|41168.101786|3458120.550000|435.384821|-0.779777|-0.827645  |189559.942118|189007.221286|0.942969   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_11|84   |      |40399.020000|41644.670000|41166.412857|3457978.680000|437.692727|-0.783717|-0.826623  |191574.923679|189987.754281|0.942863   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_12|84   |      |40393.470000|41644.670000|41164.722143|3457836.660000|440.005173|-0.787581|-0.825682  |193604.551937|190978.552570|0.942799   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_13|84   |      |40387.950000|41644.670000|41163.023929|3457694.010000|442.307629|-0.791313|-0.824788  |195636.039075|191952.926073|0.942677   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_14|84   |      |40384.170000|41644.640000|41161.289286|3457548.300000|444.533554|-0.794945|-0.824109  |197610.080554|192908.854548|0.942627   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_15|84   |      |40377.500000|41642.940000|41159.540000|3457401.360000|446.744432|-0.798871|-0.823184  |199580.587452|193953.568611|0.943042   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_16|84   |      |40377.500000|41649.650000|41158.591786|3457321.710000|448.836247|-0.801993|-0.823155  |201453.976522|194832.762580|0.942902   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_17|84   |      |40372.740000|41649.650000|41157.244048|3457208.500000|450.941890|-0.805715|-0.822129  |203348.588011|195713.317072|0.942741   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_18|84   |      |40367.070000|41649.650000|41155.893214|3457095.030000|453.055898|-0.809410|-0.821051  |205259.646501|196591.378377|0.942551   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_19|84   |      |40361.410000|41649.650000|41154.539405|3456981.310000|455.177914|-0.813077|-0.819927  |207186.933399|197468.926268|0.942345   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|_lag_20|84   |      |40355.810000|41649.650000|41153.163929|3456865.770000|457.292579|-0.816738|-0.818749  |209116.503249|198335.759559|0.942105   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
|eid    |84   |      |3           |8           |5.500000    |462           |2.073935  |0.000000 |159.618834 |4.301205     |770.108253   |0.806583   |
+-------+-----+------+------------+------------+------------+--------------+----------+---------+-----------+-------------+-------------+-----------+
```

10. Creating stream to predicting data usage value for the next day and on the 7th

We will be predicting for aggregates with used data capacity above 25%. First we will create a filter stream containing all aggregates with used data above 25% then we will create refr attributes on this stream and deploy the model for prediction.

Step 1. Creating a filter stream for aggregate above 25% then deploying the model on it.

Adding a filter :

```
“fltr”:[
  {
  "name": "aggr_above_25",
  "fqry": {
  "name": "{\"query\":
[{\"key\":\"useddatapercent\",\"cmp_op\":0,\"val\":25}]
  ,\"qtype\":1}","type": 1 },
  "fatr": [ "used_data_capacity", "cluster","hapair","aggregate",
  "growthrate","date","vid"
  ],
  "ostm": "aggrAbove25"
  }
]
```

Adding filter stream to collect data after filter :

```
{
  "name": "aggrAbove25",
  "type": 2, "inpt":["data_stream"],
  "attr": [
  { "name": "cluster","type": 5,"kysz": 64,"sidx": 1,"stat":2 },
  { "name": "hapair","type": 5,"kysz": 64,"sidx": 1,"stat": 2 },
  { "name": "aggregate","type": 5,"kysz": 64,"sidx": 1, "stat": 2},
  { "name": "useddatacapacity","type": 11, "stat": 3},
  { "name": "growthrate","type": 11 },
  { "name": "useddatapercent","type": 11},
  { "name": "vid","type": 5},
  { "name": "date","type": 5}
  ],
  "swsz": 86400
}
```

Step 3: For prediction using a model we just have to add a catr with operation "PRED"

For prediction using a model we just have to add a catr with operation "PRED".
Adding a catr :--Steps

```
1. What would you like to add (press Enter when done) [ attr (1) | catr (2) | refr (3) |
   gpby (4) | fltr (5) | join (6) | entity (7) | cep (8) | notifs (9) ]: 2
   [2 for catr]
       add computed attributes (catr)...
       attribute name (press Enter to end): nextday_usage
       [ name of the pred attribute]
  attribute type (press Enter for default (5)) [ string(5) | long(9) | double (11) ]: 11
  [type for the prediction attribute]
```

```
2.  enter the name of the intended operation from the above default ops (press Enter
    to end): PRED
    [Select the catr operation]
    enter the name of the ML model: netapp_1
    enter the name of the algo [ SVM | KMEANS | LIN | PY | IE | IE_SENT | IE_NER ]:
    [ All time series models are SVM models, So select SVM]
    enter the attribute type for the model [ NUM | STR | HYB ]: NUM
    [training attribute type]
what is the expected data format for the prediction [ LIBSVM (0) | CSV (1) | JSON (3) ] (press Enter for
default 3): 0
[ For all SVM models, the date format is 0]
```

Enter the attributes needed for prediction, we need only useddatacapacity

```
enter the input attributes on which this ops will be performed, (press Enter once done): useddatacapacity
enter the input attributes on which this ops will be performed, (press Enter once done):

should add, replace or add only if present [ add (1) | replace (2) | add only if not present (3) ]: 1
```

That's all..

Added catr in schema :

```
{
  "iatr" : ["useddatacapacity"],
  "exp_fmt" : "JSON",
  "fnr" : 1,
  "opnm" : "PRED",
  "algo" : "SVM",
  "type" : 11,
  "model" : "netapp_1",
  "attr_type" : "NUM",
  "name" : "nextdayusage"
}
```

Same for the second model. Add a catr for prediction.

11. Sending alert for aggregates who will be reaching above 90% usage on 7th day

Adding a catr to calculate percentage :

```
"catr" : [
  {
  "name" : "7dayusedper",
  "opnm" : "PERCENT",
  "iatr" : [
  "7dayusage",
  "totaldatacapacity"
  ],
  "type" : 11
  }
]
```

Adding a cep query to stream :

```
"cepq": [
  {
  "name": "7dayusage",
  "type": 6,
  "tloc": 1000,
  "fqry": {
  "type": 1,
  "name":
  "{\"query\":
[{\"key\":\"7dayusedper\",\"cmp_op\":0,\"val\":\"90\"}],\"qtype\":1}"
  },
  "notf": 807
  }
]
```

12. Monitoring next data volume usage

Adding group by :

```
"gpby": [
  { "iatr": "nextdayusage",
  "gpat": ["cluster",”hapair”,”aggregate” ],
  "kysz": 124, "gran": 3600, "stat": 3
  }
]
```

Query to get result

```
bangdb> select aggr(nextdayusage) from netapp.aggrAbove25 groupby
cluster:hapair:aggregate
```

CONCLUSION

In this use-case, we have shown some features of the platform. It's query structure is very simple and easy to understand.For more information please visit bangdb.com.

We deployed the BangDB server along with frontend dashboards for visualization, reporting and further configuring the solution.

BangDB runs the solution in autonomous mode and sends notifications and reports as required.