NSF/IUCRC CAC PROJECT

MONITORING, VISUALIZING, AND PREDICTING HEALTH STATUS OF HPC CENTERS

Jie Li

PhD Student, TTU

11/08/2019

Advisors:

Mr. Jon Hass, SW Architect, Dell Inc.

Dr. Alan Sill, Managing Director, HPCC, TTU

Dr. Yong Chen, Associate Professor, CS Dept, TTU

Dr. Tommy Dang, Assistant Professor, CS Dept, TTU

Receive requests from analytics client (HiperViz)

time range,
e.g. 2019-04-20T00:00:00Z, 2019-04-21T00:00:00Z
time interval, e.g. 30m
data type, e.g. Max, Min, Average

Generate corresponding influxDB query requests

SELECT CPU_Usage FROM CPU_Usage WHERE host='10.101.1.1' AND time >= '2019-04-20T00:00:00Z' AND time <= '2019-04-21T00:00:00Z' GROUP BY time(30m) SLIMIT 1
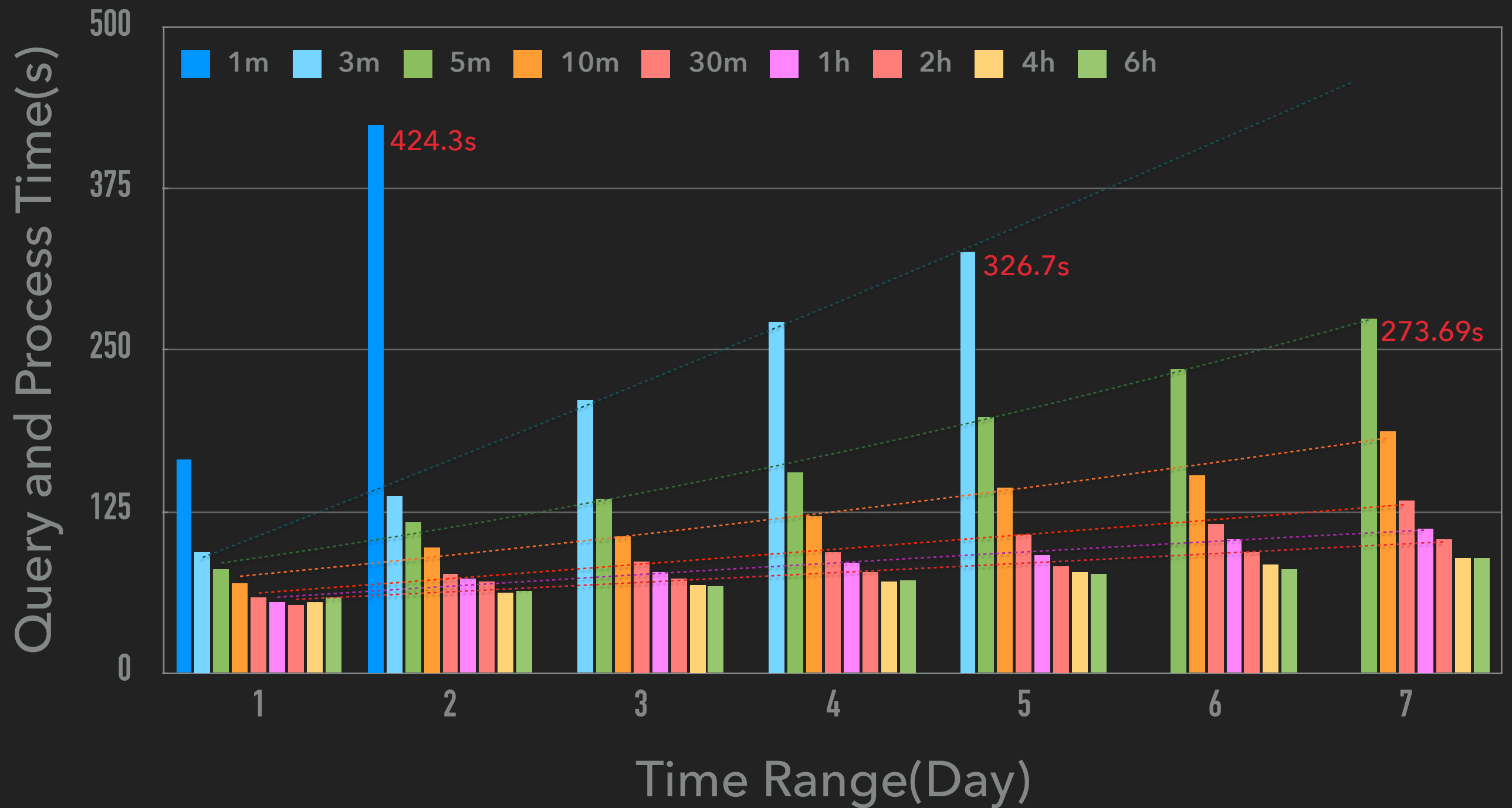
Process data returned from influxDB;
Convert it to csv dataframe;
Return dataframe to analytics client

‣ Move data processing from front end to back end
‣ Provide a uniform API to analytics client
‣ Act as a middleware that deals with different database design

Time Occupation of querying ONE host info

Profiling tool: cProfile
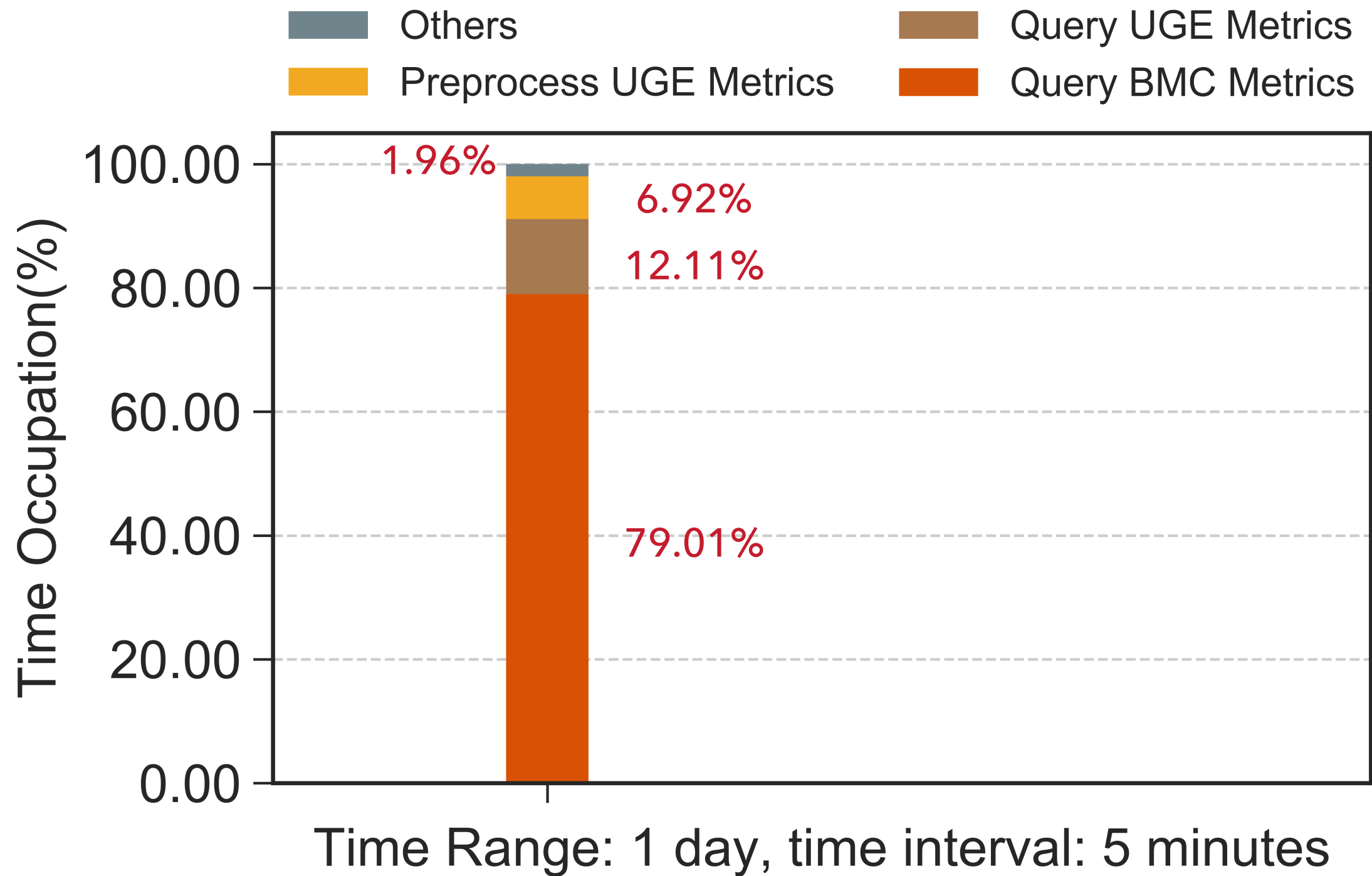
Query and process time of MetricsBuilder API

QUESTIONS?/COMMENTS?