SOLUTIONS ENGINEER TECHNICAL EXERCISE

Chris Reites

October 2018



BEFORE WE GET STARTED...





Here's a little bit about me. My name is Chris Reites, and I'm currently in presales at CA Technologies helping to sell their DevOps/Continuous Testing software solutions. Prior to joining CA, I spent 16 years at FedEx as a software developer, QA Engineer, and Enterprise Architect. Having spent a lot of time as a customer of enterprise software, I believe I have great perspective and appreciation for the challenges that my customers face. The reason why I made the move into presales, and why I enjoy it so much, is that I love being able to combine the excitement of sales with my passion for helping customers solve their problems and bring more value into their own organizations.

Now I'm looking for my next adventure. That's where you come in! I've been talking to people I know at Datadog, and it sounds like an amazing place to be. Even before I knew about this technical exercise, I signed up for a free trial at datadoghq.com and started playing around with the technology. I was highly impressed and that's what ultimately convinced me that I wanted to join your team. My hope is that after reviewing my submission, you'll feel the same about me.

When I'm not busy helping customers with their digital transformations, I enjoy living in the Tampa area with my wife and two children. On the weekends you can usually find us boating or at the beach. I also like running, but to ensure I don't become *too* healthy, I also enjoy brewing my own beer.

https://www.linkedin.com/in/chris-reites/



MY ENVIRONMENT

For this exercise, I chose to use an AWS instance I already had setup. It's running Ubuntu 16.04 and I have a few java apps running **ubuntu**. there that I was able to play around with to check out more Datadog functionality. I previously had configured my applications to use SQL Server as the backend database, but because the current version of that integration doesn't work on Linux, I installed MySQL and changed my apps to use that for the backend. Finally, in order to create some interesting moments that I could look at metrics for, I used BlazeMeter, which is a cloud-based load generation solution, to create a lot of traffic to my applications.











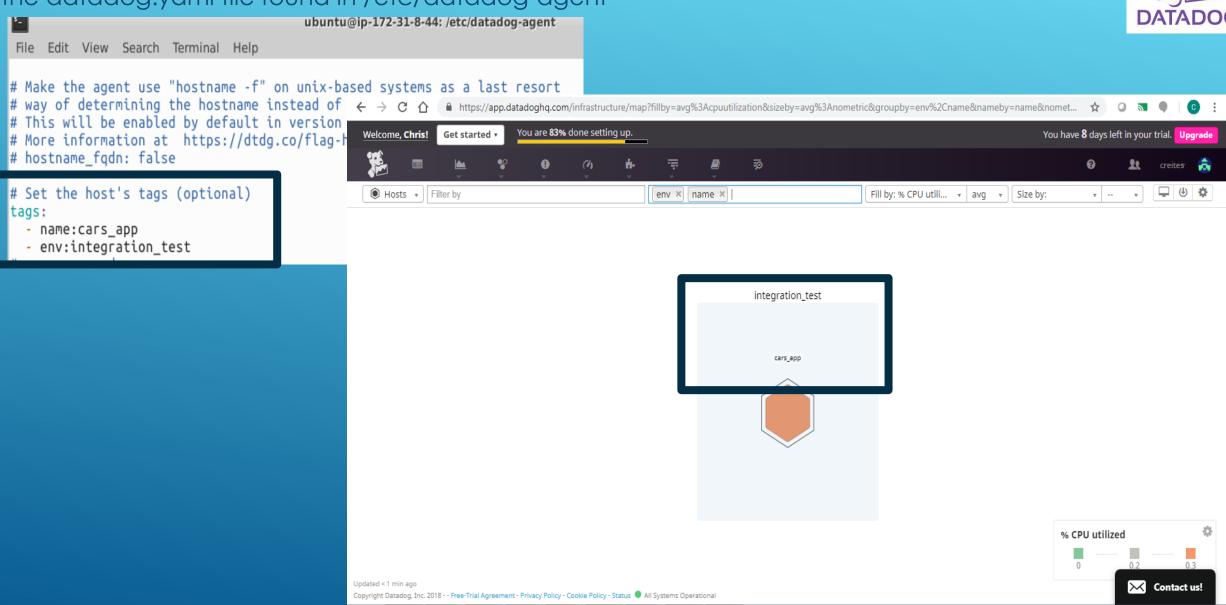




COLLECTING METRICS

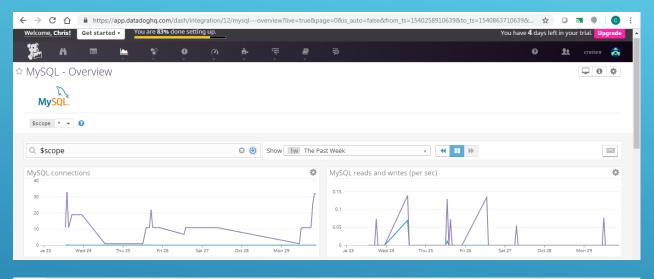
The first thing I needed to do was get the Datadog Agent stood up on my server, and configure it with a meaningful name and some helpful tags. To do this, I simply had to edit the datadog.yaml file found in /etc/datadog-agent

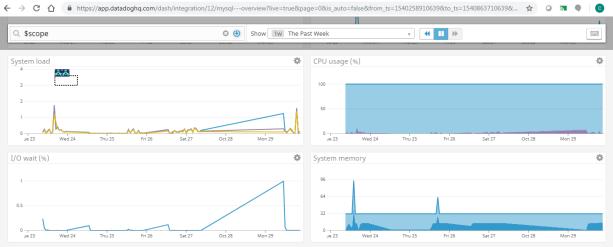


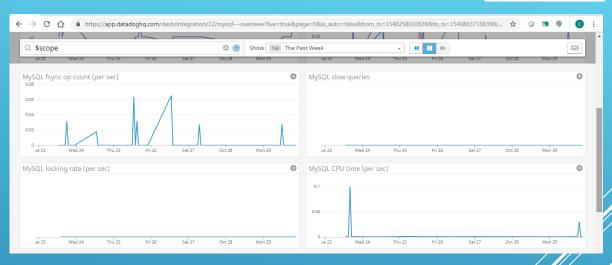


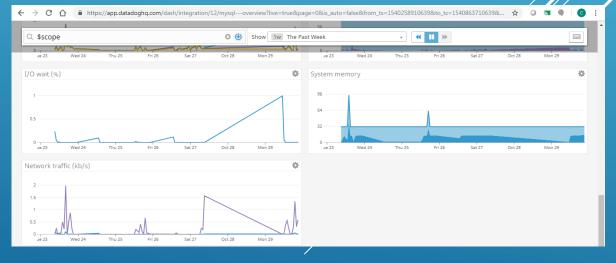
Next I added in the MySQL integration so that I could keep an eye on my database metrics. One of the nice features of Datadog is that when you install integrations, you get default Dashboards which include popular metrics. Here are some shots of my default MySQL Dashboard. It's pretty easy to spot when I was running my load tests!





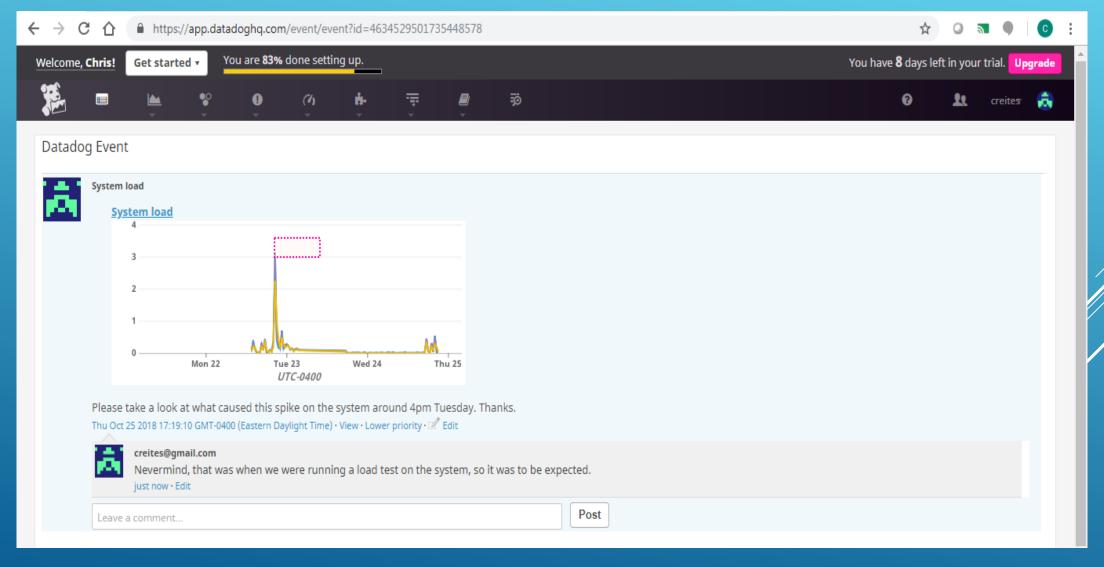






Speaking of being easy to spot... When you do see something that catches your eye, it's very easy to collaborate with your team. Here, I created a snapshot of one of the spikes, annotated it, and then sent if off for the team to look into. Unfortunately, I'm the only one on my team, so it was up to me to figure out what was going on and reply. Sometimes it feels like I have to do everything!



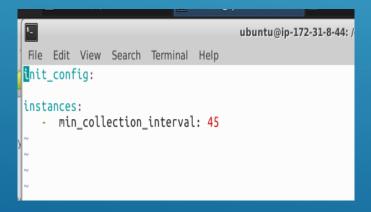


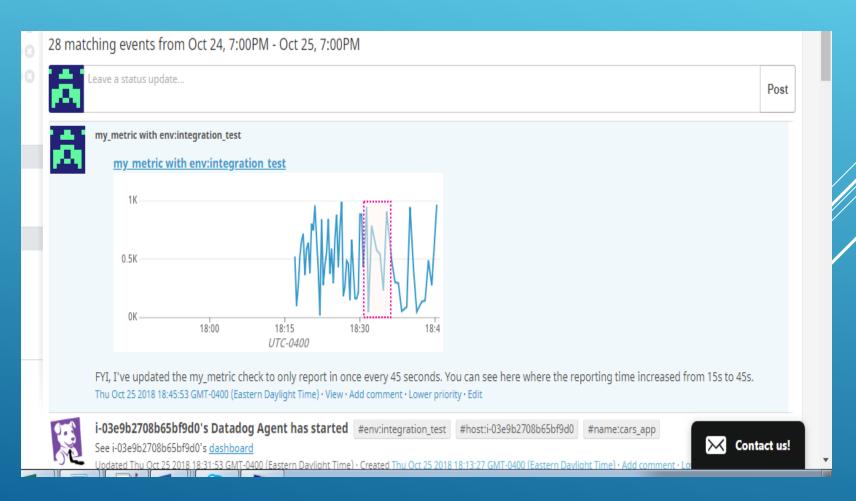
In order to further test out the metric gathering capabilities of Datadog, I wrote a simple python script that generates a random number between 1 and 1000 and then submits the value to the Agent with the name my_metric. After setting up the script, I then changed the collection interval to be every 45 seconds instead of the default 15 seconds.



```
ubuntu@ip-172-31-8-44: /er
File Edit View Search Terminal Help
__version__="1.0.0"
import time
import random
from checks import AgentCheck

class DemoCheck(AgentCheck):
    def check(self, instance):
        random_nbr = random.randint(1,1001)
        self.gauge('my_metric', random_nbr)
        random_nbr)
```





For fun (and bonus points!) I added in the ability to dynamically change the collection interval by simply updating a delay file on the filesystem. Whenever the check is fired, it will re-read the delay file and skip reporting the metric if the collection interval hasn't passed.

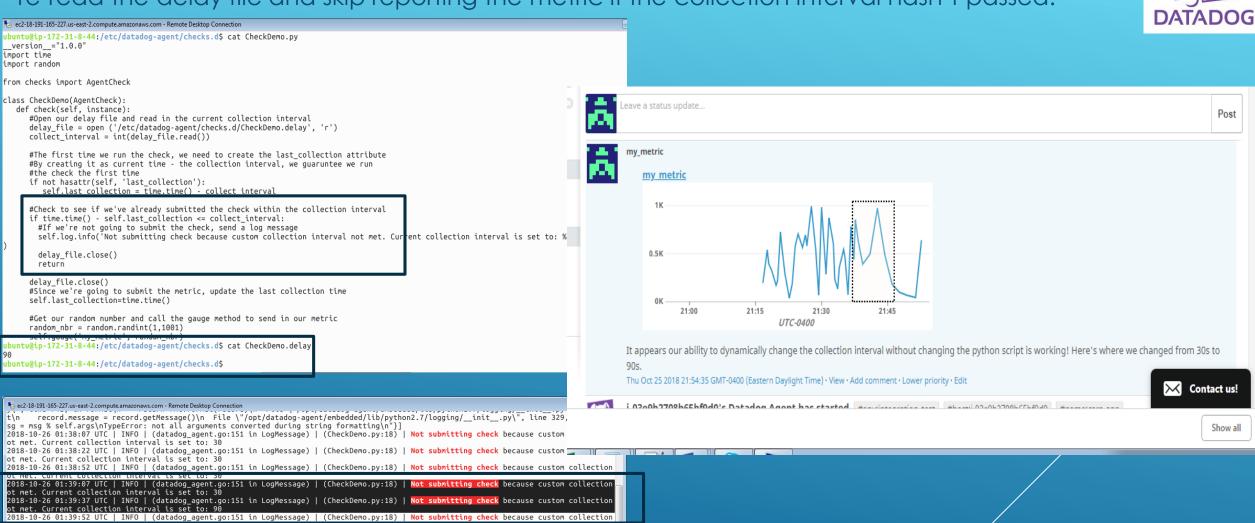
2018-10-26 01:40:07 UTC | INFO | (datadog_agent.go:151 in LogMessage) | (CheckDemo.py:18) | Not submitting check because custom collection

2018-10-26 01:40:22 UTC | INFO | (datadog_agent.go:151 in LogMessage) | (CheckDemo.py:18) | Not submitting check because custom collection

ot met. Current collection interval is set to: 90

ot met. Current collection interval is set to: 90





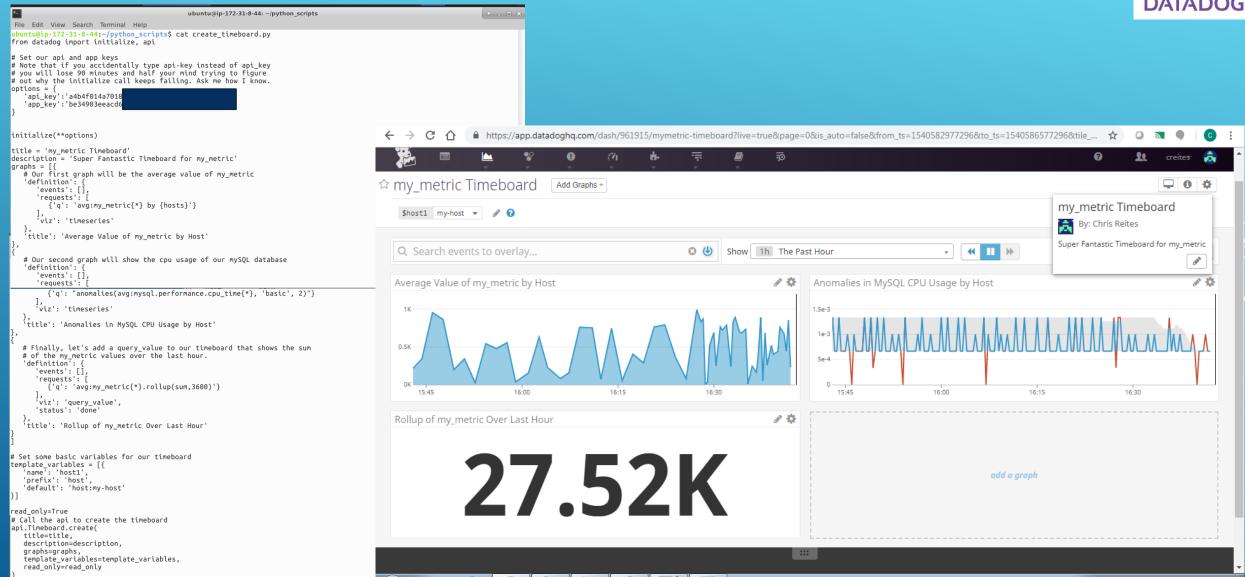


VISUALIZING DATA

The Datadog platform includes very powerful APIs that can be used to programmatically create many of the same assets you can build manually inside the web frontend. One example is Timeboards. Here is a python script I wrote to create a custom Timeboard.

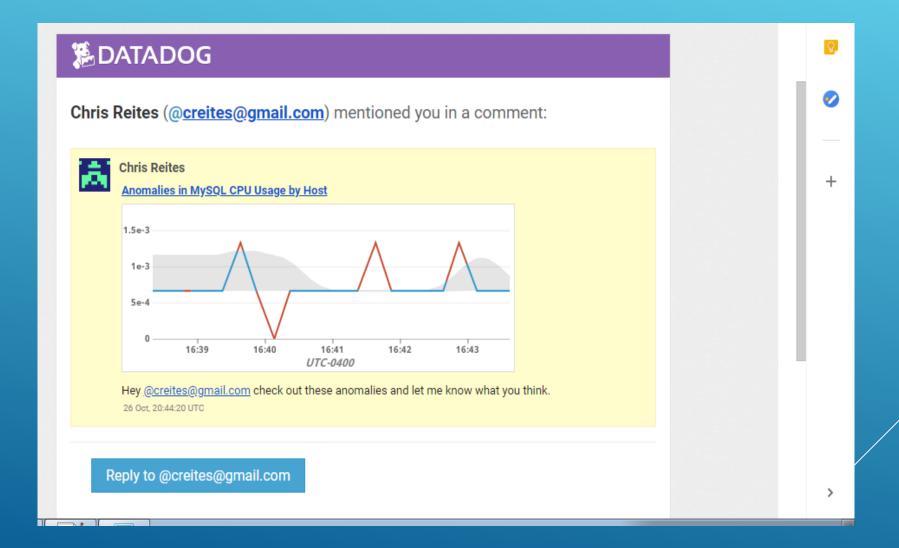
ubuntu@ip-172-31-8-44:~/python_scripts\$





Here's an email containing a snapshot that I generated from my custom Timeboard. In this case, I'm looking at a 5 minute window of my graph that shows Anomalies in the CPU usage of MySQL. The Anomalies function is powerful because it compares actual metrics to expected values derived from historical trend analysis.







MONITORING DATA

One of the most useful aspects of the Datadog platform is the ability to create monitors for anything you want to be notified about. In this example, I created a monitor to alert me with either a warning or an alert if my custom random number metric rose above an average of either 500 or 800 over a 5 minute period. I'll also be notified if my metric doesn't send in any data over the last 10 minutes.



← → C ↑ https://app.datadoghq.com/monitors#6856098/edit	☆ ② 3 ● ⓒ :
1 Choose the detection method	•
Threshold Alert Change Alert Anomaly Detection Outliers Alert Forecast Alert	
An alert is triggered whenever a metric crosses a threshold.	0
2 Define the metric	Source Edit
a Metric my_metric flom host:i-03e9b2708b65bf9d0 × excluding (none) avg by (everything)	
	Advanced
Simple Alert Trigger a single alert when your metric satisfies your alert conditions.	0
3 Set alert conditions	
Trigger when the metric is above verthe threshold on average verthe during the last 5 minutes verther sminutes	
Alert threshold: 800 (0.8K)	
Warning threshold: 500 (0.5K)	
Alert recovery threshold: Alere recovery threshold (opt	
Warning recovery threshold: Warning recovery threshold	
Require a full window of data for evaluation.	9
Note: We highly recommend you select "Do Not Require" for sparse metrics, otherwise some evaluations will be skipped.	
Note. We nightly recommend you select. Do Not Require for sparse metrics, otherwise some evaluations will be skipped.	
Notify ▼ if data is missing for more than 10 minutes.	Contact us!

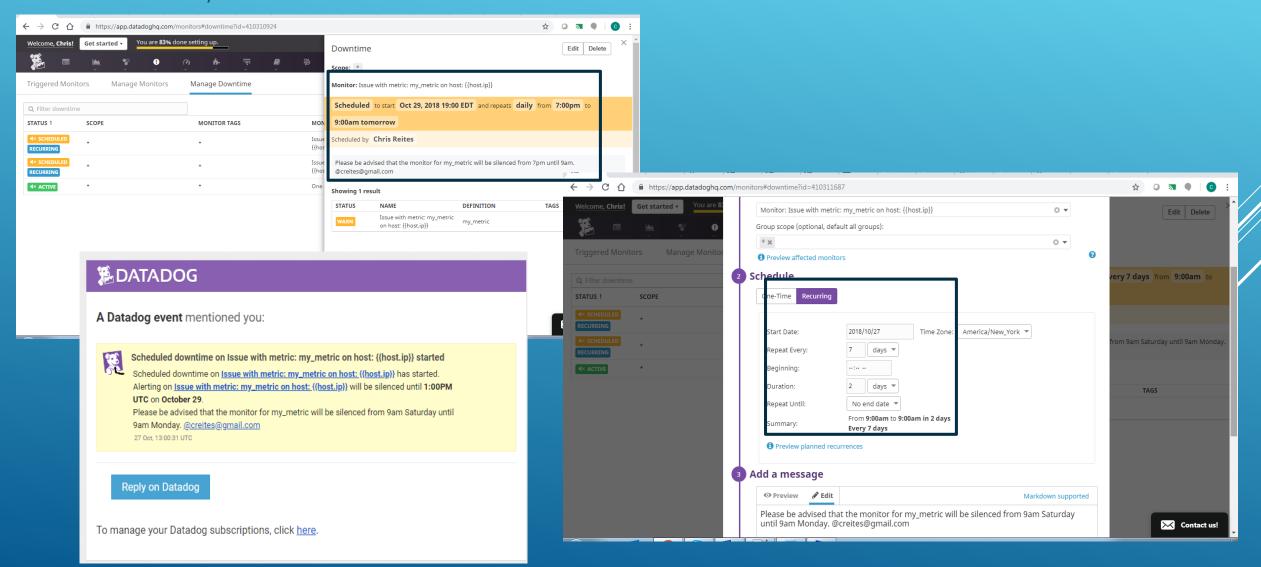
When setting up monitors, you can create highly customized messages that vary depending on things such as warnings vs alerts. You can also include variables such as the host IP address and metric values to make the monitor emails more meaningful.



← → C ↑ https://app.datadoghq.com/monitors#6856098/edit	☆ ② 🗃 🎈 ③ :	DAIADO
[Never] value automatically resolve this event from a no data state.	•	
Delay evaluation by 0 seconds		
Say what's happening	Use message template variables 🔞	
☐ Preview	Markdown supported	
{{#is_warning}} Please be advised that the average value of my_metric over the last 5 minutes has crossed above WARN level on host {{host.ip}}. Compy, metric over the last 5 minutes is: {{value}} {{/is_warning}} {{#is_alert}}		
Please be advised that the average value of my_metric over the last 5 minutes has crossed above ALERT level on host {{host.ip}}. Cumy_metric over the last 5 minutes is: {{value}} {{/is_alert}} {{#is_no_data}} Please be advised that my_metric has not reported data for the last 10 minutes on host {{host.ip}} {{/is_no_data}} @creites@gmail.com	ur rent average value c Barbara DATADOG	
Tags: Select or add related tags	[Warn] Issue with metric: my_metric on host: 172.31.8.44	
[Never] vernotify if the monitor has not been resolved.	Please be advised that the average value of my_metric over the last 5 minute	s has crossed above WARN level on
5 Notify your team [Chris Reites ×]	host 172.31.8.44. Current average value of my_metric over the last 5 minutes	s is: 547.75
Do not notify ▼ alert recipients when this alert is modified	@creites@gmail.com 1K 0.5K 0.5K	

And because I don't want to be bothered while watching football on Saturdays or Sundays, it's always great to be able to schedule downtime windows when the monitors will automagically mute themselves. Here are two that when combined will mute the monitor on weekday evenings and also Saturday starting at 9am through Monday at 9am. We can also make sure the system sends us an email when it mutes a monitor.







COLLECTING APM DATA

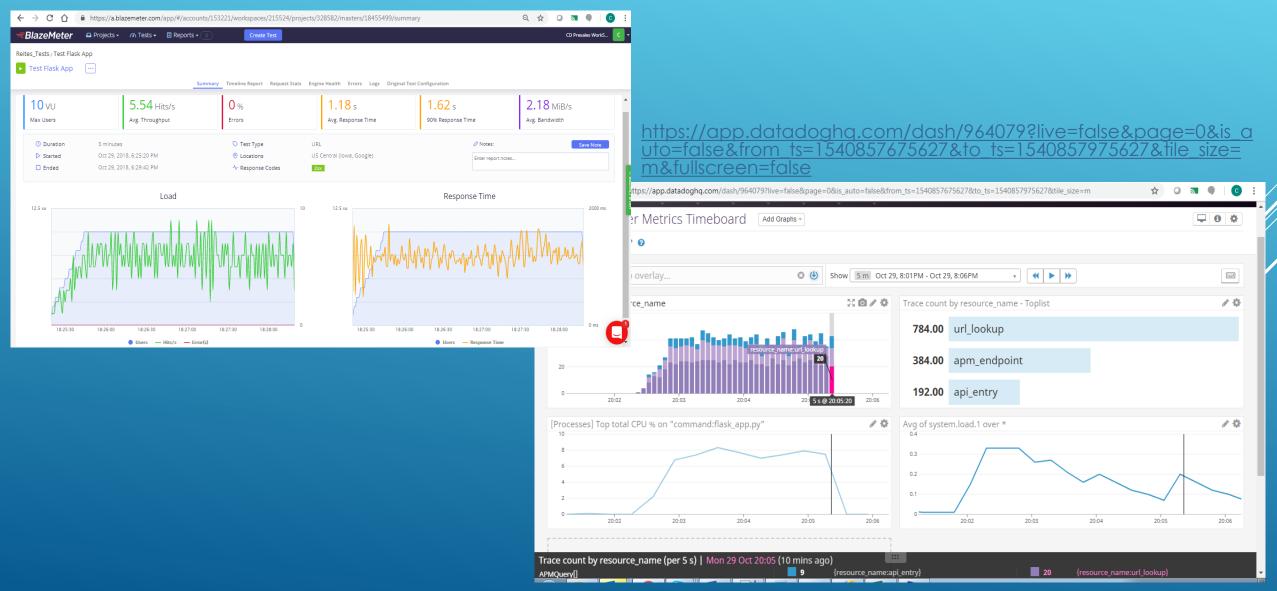
Using the Datadog platform, it's very easy to instrument applications for performance monitoring, either by running the process with ddtrace-run or by explicitly instrumenting your application. In the below example, I've instrumented a Flask application to feed metrics related to the performance of several routes.



```
ubuntu@ip-172-31-8-44: ~/flask scripts
File Edit View Search Terminal Help
ubuntu@ip-172-31-8-44:~/flask scripts$
 buntu@ip-172-31-8-44:~/flask scripts$ cat flask app.pv
#!/usr/bin/env python
from flask import Flask
from flask import request
from time import sleep
#import the necessarry modules for tracing our app
import blinker as
from ddtrace import tracer
from ddtrace.contrib.flask import TraceMiddleware
import sys
import random
import requests
# Have flask use stdout as the logger
main_logger = logging.getLogger()
main_logger.setLevel(logging.DEBUG)
c = logging.StreamHandler(sys.stdout)
formatter = logging.Formatter('%(asctime)s - %(name)s - %(levelname)s - %(message)s')
c.setFormatter(formatter)
main_logger.addHandler(c)
app = Flask( name )
# Instrument our app
traced app = TraceMiddleware(app, tracer, service="sample-flask-app", distributed tracing=False)
def api entry():
   return 'Entrypoint to the Application'
# Create a new method that does something interesting that we can measure
@app.route('/url_lookup')
def url_lookup():
   url_name= request.args.get("url-name")
   full_url= "http://" + url_name
    # Trace how much time this portion of the method takes
   with tracer.trace("setup_overhead", service = "url-lookup"):
       # Artificially insert some random overhead time to make things interesting
        random_nbr = float(random.randint(250,750))
        sleep (float(random nbr/1000))
    # Trace how much time the actual URL call takes
   with tracer.trace("call_url") as span:
       # By setting a tag in the span, we can see which URL was being called when we look at our metrics
        span.set_tag("url-name", url_name)
        return requests.get(full_url).content
@app.route('/api/apm')
def apm endpoint():
   return 'Getting APM Started'
@app.route('/api/trace')
def trace_endpoint():
   return 'Posting Traces
if __name__ == '__main__':
   app.run(host='0.0.0.0', port='5050')
 buntu@ip-172-31-8-44:~/flask_scripts$
  ountu@ip-172-31-8-44:~/flask_scripts$
```

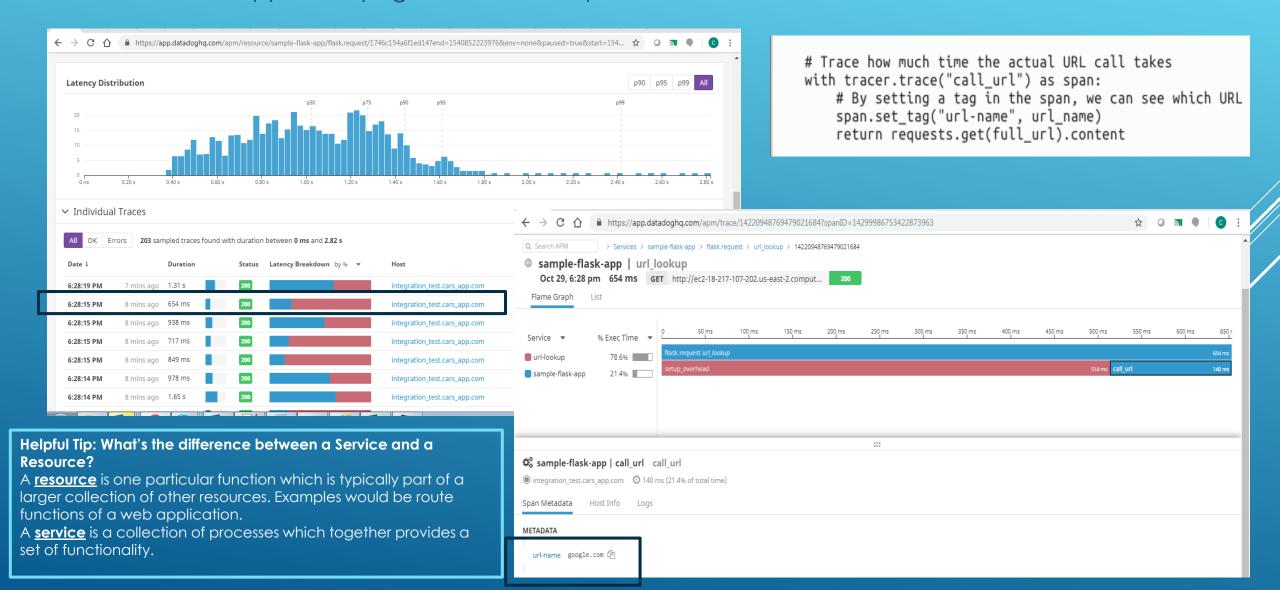
After I instrumented my application, I then used BlazeMeter, a cloud based load generation tool, in order to send thousands of requests to my application. The Timeboard below shows the counts by resource name, along with metrics about the CPU usage of the python app, and overall CPU utilization.





If we dive in a little deeper into our APM metrics, we can analyze latency as well. Because we setup a span with a tag name that matches the URL being loaded, we can also drill into individual traces and see the breakdown of time spent, along with the span data which shows the URL our app was trying to load on this particular call.







BUT WAIT... THERE'S MORE!!!

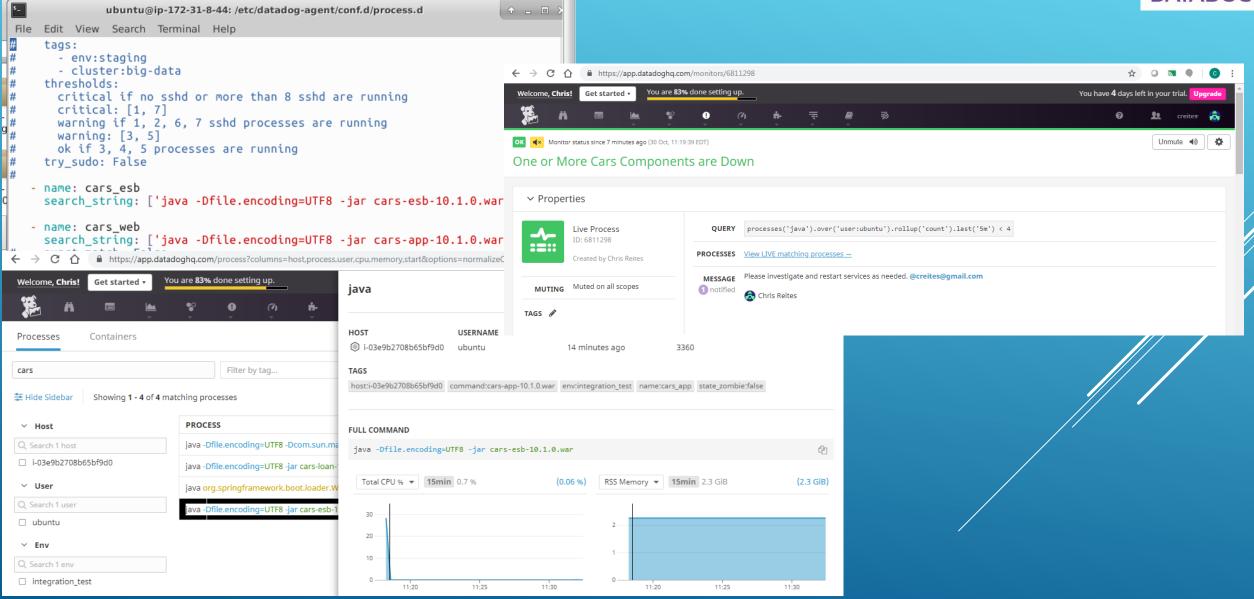
Prior to beginning this exercise, I had signed up for a free account on Datadog and was exploring functionality on my own. Because I had java applications, I installed the Java integration and ran some tests. The first thing I had to do was setup my application to accept JMX connections. I then had to configure the conf.yaml for the jmx integration to know how to connect. After that it was time to run some tests using BlazeMeter. I was most interested in two aspects of my application: The Garbage Collection metrics as well as the backend database performance.



```
if [ "$WEB" = "true" ]; then
           echo "Starting WEB'
           java -Dfile.encoding=UTF8 -Dcom.sun.management.jmxremote.port=7199 -Dcom.sun.management.jmxremote.ssl=false $WEB AG
ENT -jar cars-app-10.1.0.war &
 instances:
  host: localhost
  port: 7199
                                                                                                                                             C 🏠 https://app.datadoghq.com/dash/957160/chriss-timeboard-23-oct-2018-1054?live=false&page=0&is auto=false&from ts=1540855121731&to ts=15408... 🏠
  user: monitorRole
  password: DD READONLY JMX
                                                                                                                                        Q Search events to overlay..
                                                                                                                                                                                               3 Show 6 m Oct 29, 7:18PM - Oct 29, 7:24PM
                                                                                                                                                                                                                                              → | (4 | ▶ | )>
                                                                                                                                                                                                               Avg of mysql.performance.com_select over cars_inv_db
                                                                                                                                        [Processes] Top total CPU %
  BlazeMeter Projects - A Tests - Reports -
Reites Tests / Cars App Stress Test
Cars App Stress Test ...
   200 vu
                                                                                                            20.87 MiB/s
                        405.76 Hits/s
                                                                  301.02 \, \text{ms}
                                             Errors
                                                                  Avg. Response Time
   Max Users
                        Avg. Throughput
                                                                                       90% Response Time
                                                                                                                                                                                                                      Avg of jvm.gc.cms.count over name:cars_app
                                                                                                                                        Avg of jvm.gc.parnew.time over name:cars_app
                 Oct 29 2018 7:23:37 PM
    ☐ Ended
                                                                                           Response Time
                                                                                                                                                                                                                                                                                             19:24
                                                                                                                                        Avg of mysgl.performance.open_tables over cars inv db
                                                                                                                                                                                                                       Avg of system.load.1 over *
                                                                                                                                                                                                                                                                                             1 0
```

For this same application, I set up a monitor to make sure all 4 components of the application were always up and running. This also allowed me to play around with enabling the process check functionality.







FINAL QUESTION

IMPROVING THE EFFICIENCY OF A BIKE SHARING BUSINESS WITH DATADOG



After having used several of the features that make up the Datadog platform, I began thinking of a fun use case that Datadog could bring value to. Having lived a big portion of my life in Central Florida (Mickey Mouse Town), a lot of the things I thought about were theme park related (ride wait times, parking, trend analysis of concession sales, etc...) but then while riding my bike to lunch today, I thought of a great use case having to do with bike shares.

Bike share programs are a great solution to commuting woes in large cities, however they're often plagued by a common problem. As users rent bikes and transport them to wherever they're headed, you end up with too many bikes in some places, and not enough bikes where you need them.

By utilizing the Datadog platform, you could keep track of how many bikes were currently in each location, and based on monitoring criteria, you could notify the appropriate "bike relocation" staff to transport bikes from areas with low demand to areas with high demand. You could also use historical trend analysis to proactively move bikes into areas where you are forecasting upcoming higher demand.







I really enjoyed getting to explore the Datadog platform over the past week. Datadog has a fantastic product from what I've seen. So many customers I talk with today are using several different tools from multiple vendors to accomplish what Datadog can do on a single cloud based platform. I would greatly value the chance to join the team there and help drive the Datadog business forward while enabling customers to solve the many challenges they have around infrastructure, application, and log monitoring. If you have any questions for me about this exercise, please feel free to reach out to me at creites@gmail.com.

Thank You, Chris