What to expect from this session

- Brief review of serverless (API Gateway/Lambda)
- Customizing your API
  - Custom Timeouts
  - Gateway Responses
- Monitoring Your API
  - Amazon CloudWatch Metrics/Alarms
  - Amazon CloudWatch Logs
- Protecting Your API
  - Throttling
  - Authorization
  - Usage Plans
- Updating Your API
  - Canary Release Deployments

Serverless means…

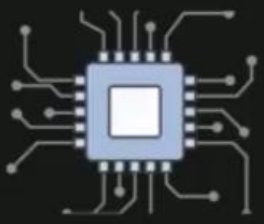
No servers to provision or manage

Scales with usage

Never pay for idle

Availability and fault tolerance built in

Amazon API Gateway

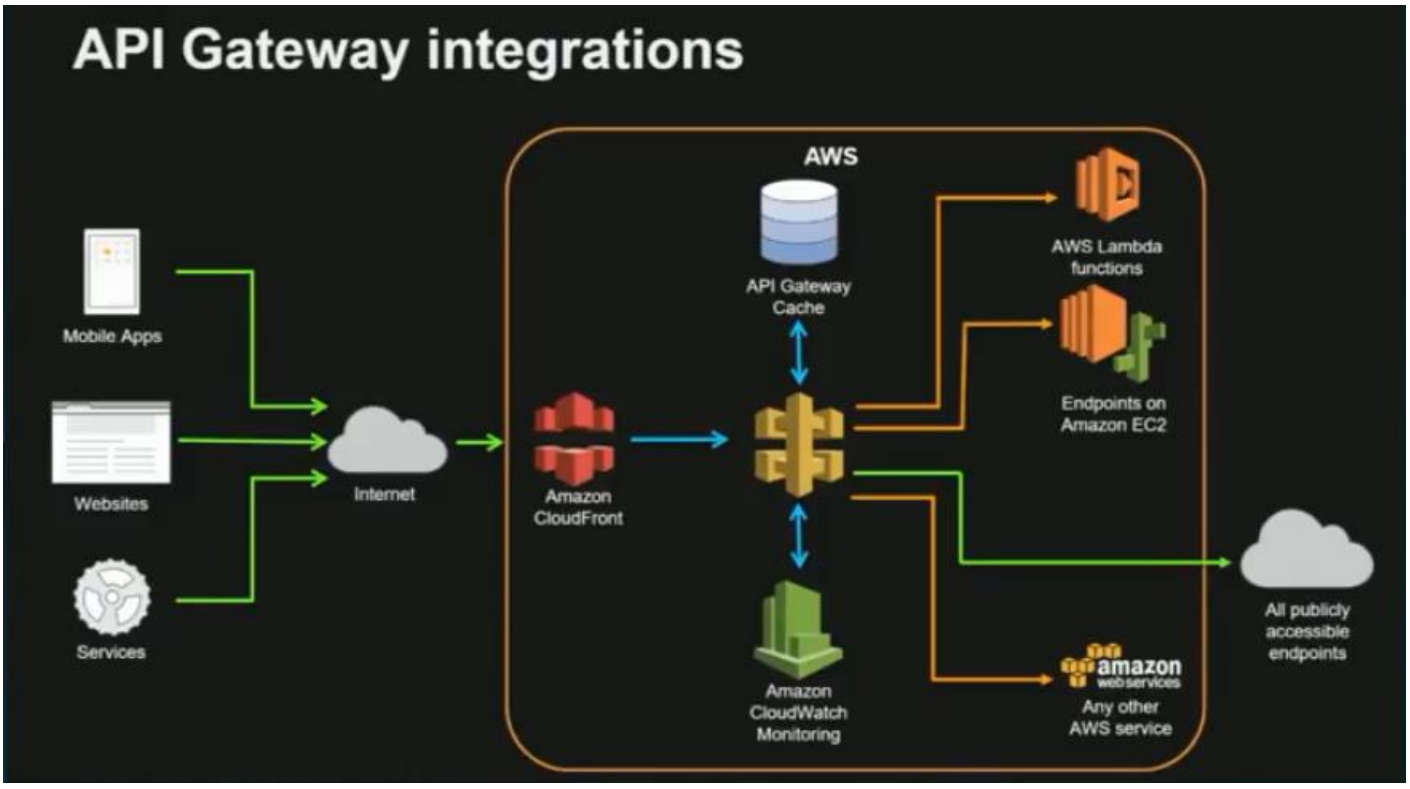


Create a unified API frontend for multiple micro-services

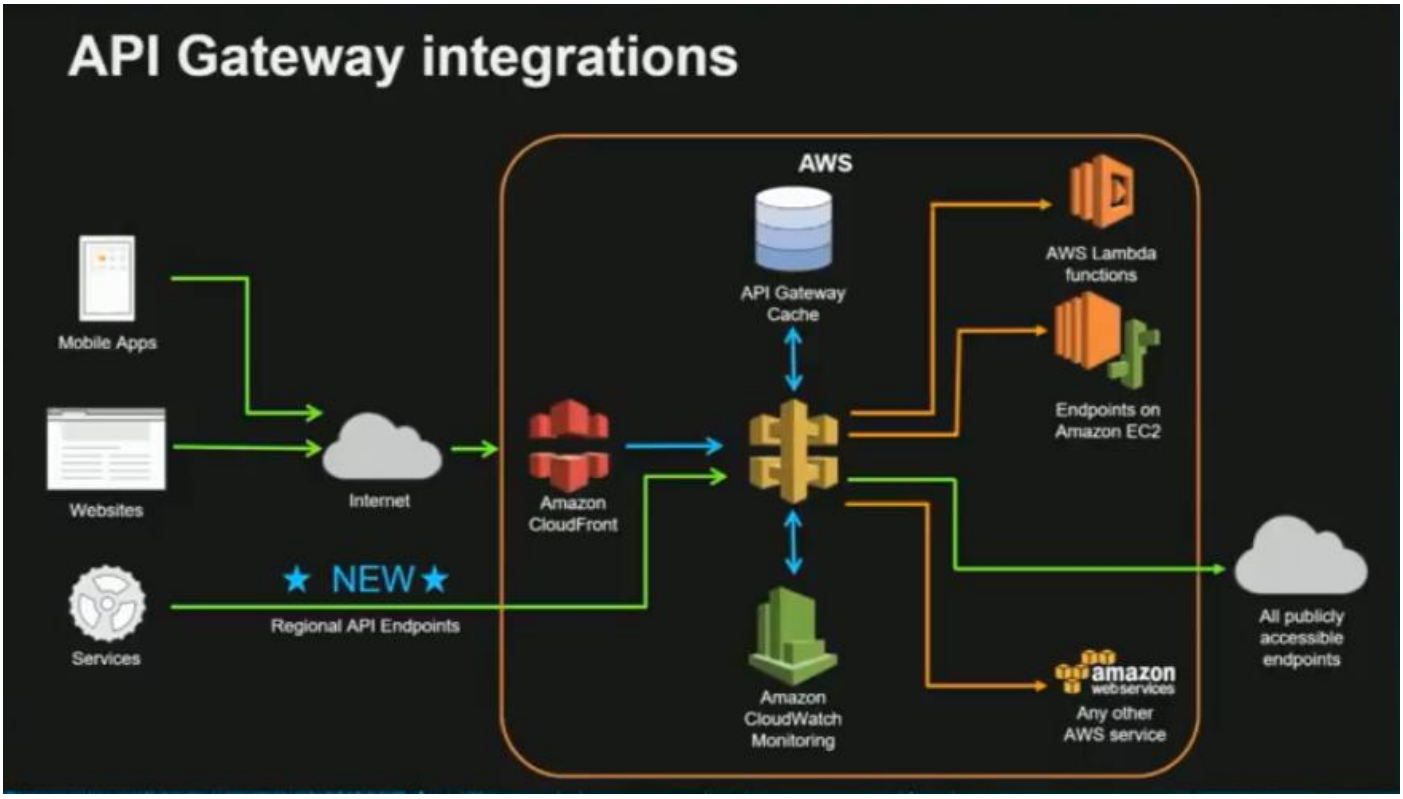
DDoS protection and throttling for your backend

Authenticate and authorize requests to a backend

Throttle, meter, and monetize API usage by third-party developers



We can front our API Gateway with CloudFront as above



You can also remove CloudFront from the loop if desired as above

AWS Lambda

Bring your own code

- Node.js, Java, Python, C#
- Bring your own libraries (even native ones)

Simple resource model

- Select power rating from 128 MB to 1.5 GB
- CPU and network allocated proportionately

Flexible use

- Synchronous or asynchronous
- Integrated with other AWS services

Flexible authorization

- Securely grant access to resources and VPCs
- Fine-grained control for invoking your functions

Meet Doug

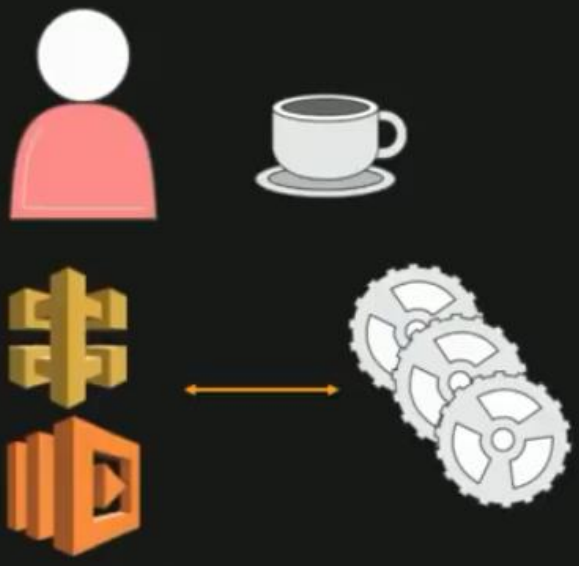
Doug loves coffee.

Doug also writes apps.

Doug built TAMPR—a service for sharing reviews of coffee and coffee shops.

Doug built the TAMPR backend serverless, with API Gateway and AWS Lambda.

External data sources



Doug built the first version of TAMPR in part by aggregating data from other sources. The service will fetch some data on demand.

Unfortunately, the service is not as reliable as he'd like, and sometimes requests timeout after 30 seconds.

Doug would like to be able to fail faste in these cases so his clients can retry.

Customizing Your API: Custom Timeouts

Custom timeouts

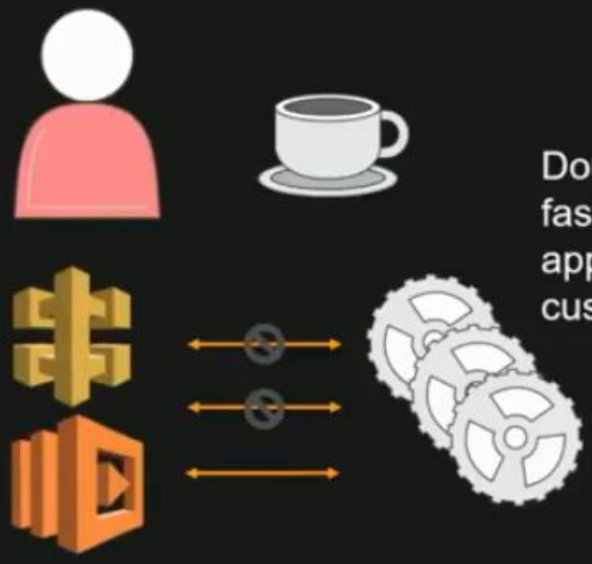
AWS Lambda

- Set your maximum invocation time
- Set per function

API Gateway★ NEW★

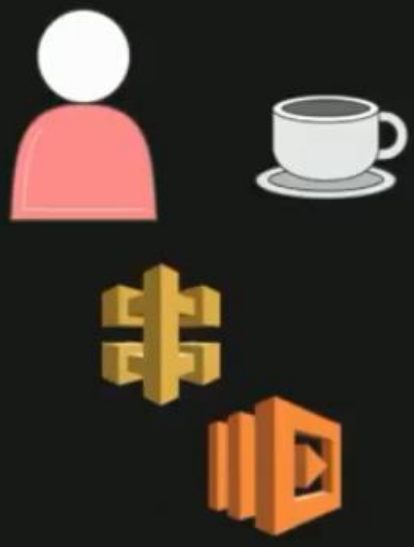
- Set your integration timeout (1–29 s)
- Works with any integration type (HTTP/Lambda/AWS)
- Set per method

Check in with Doug



Doug now can configure his API to fail fast and build retry logic in his application for a more reliable customer experience.

TAMPR website



Doug has already built native mobile applications, but wants to add a website with at least a subset of the functionality before going live with TAMPR.

He's gotten CORS working in the happy path, but a lot of his error handling is failing due to lack of CORS headers.

He needs a way to customize the erro responses coming from API Gateway to add CORS headers.

Customizing Your API: Gateway Responses

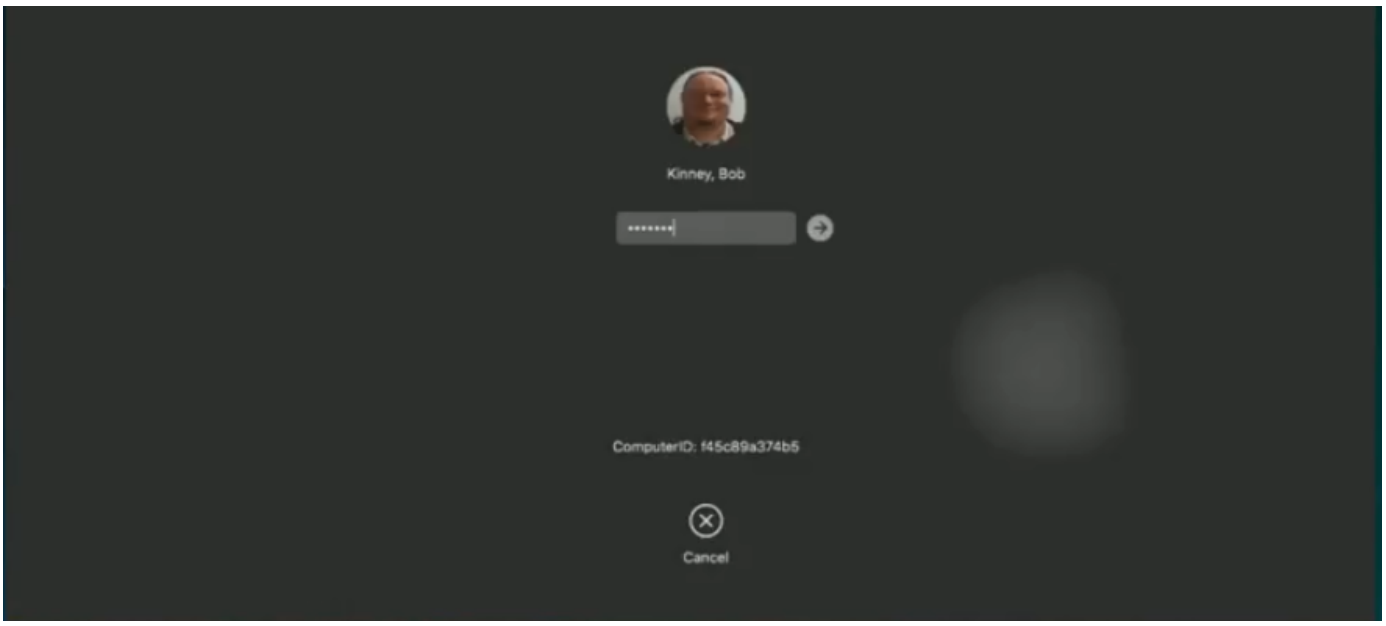
Gateway responses

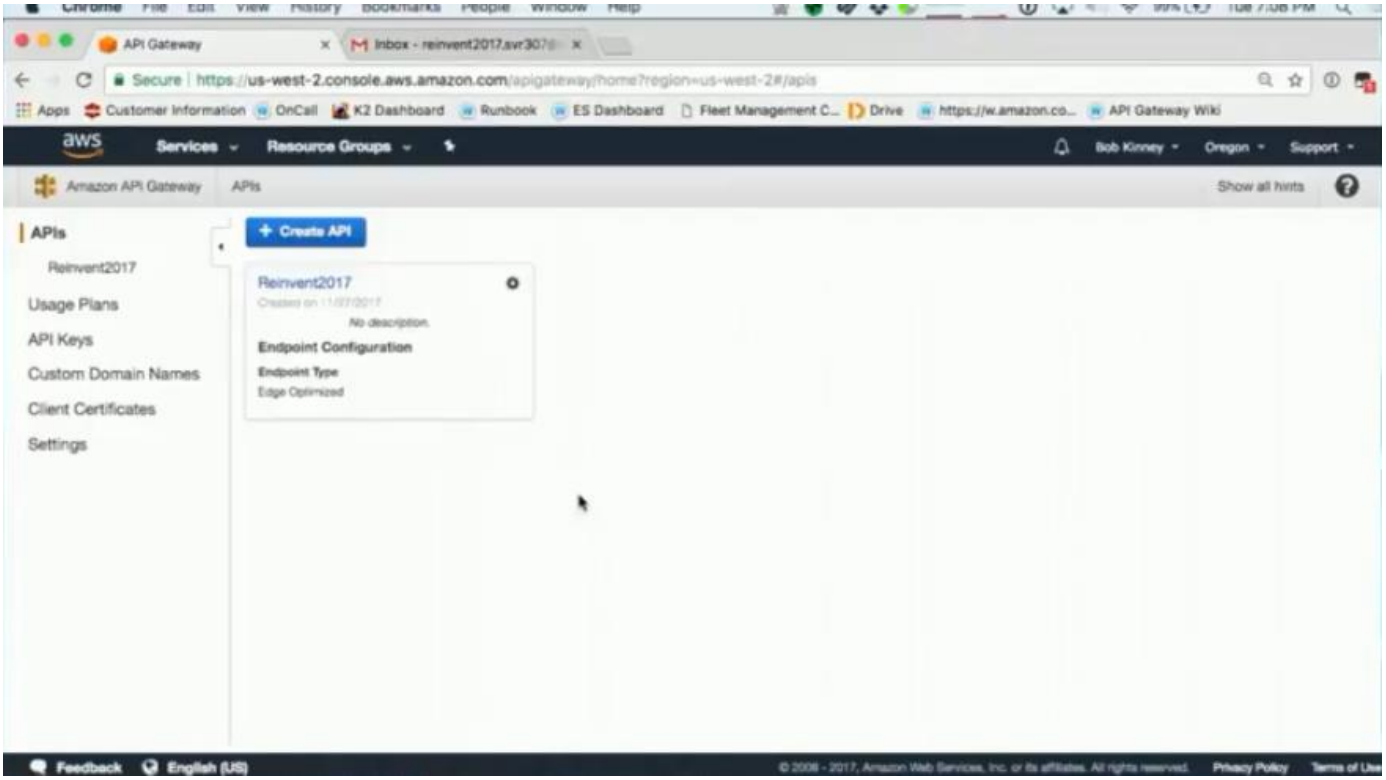
Allows customization of various error responses

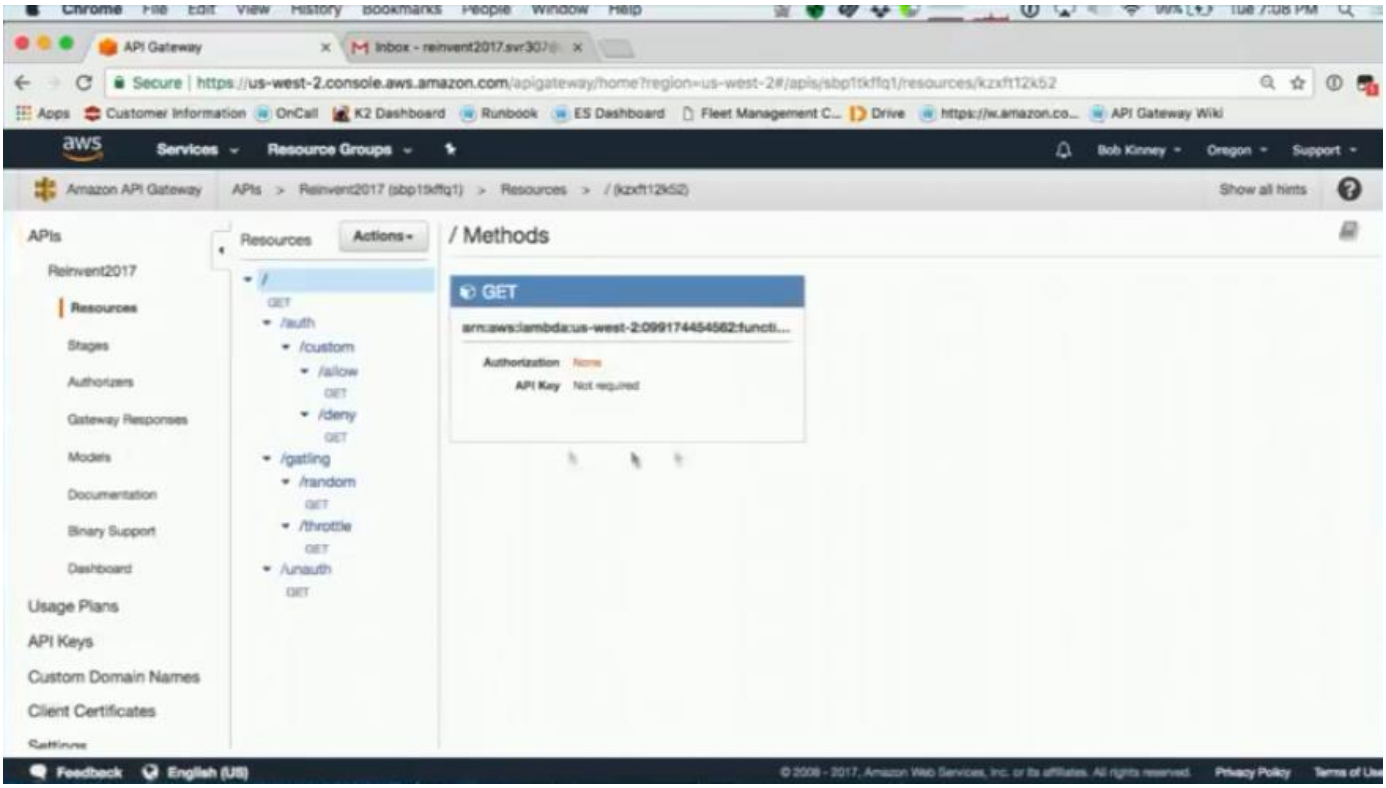
- Change HTTP status code
- Modify body content
- Add headers

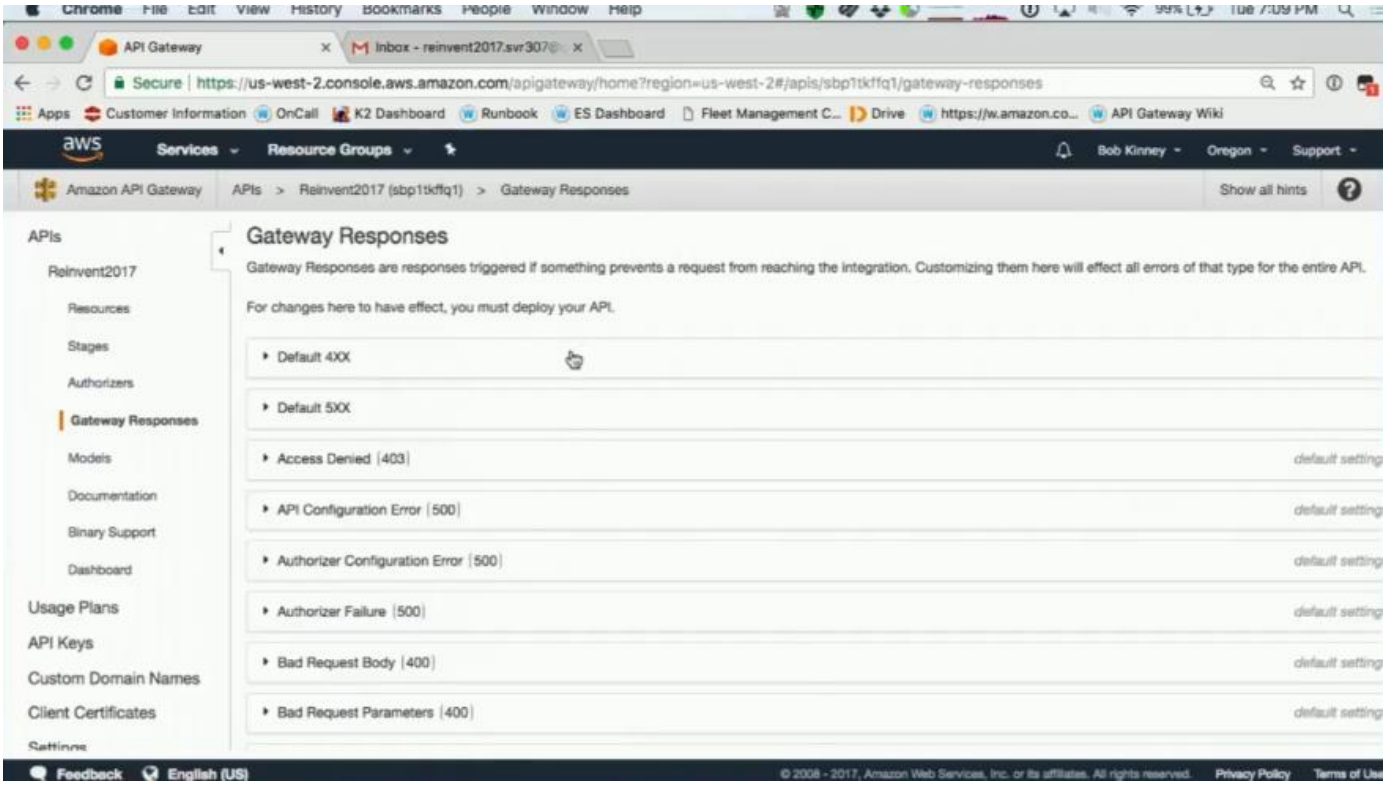
Can customize specific responses and/or modify default 4XX/5XX

Demo

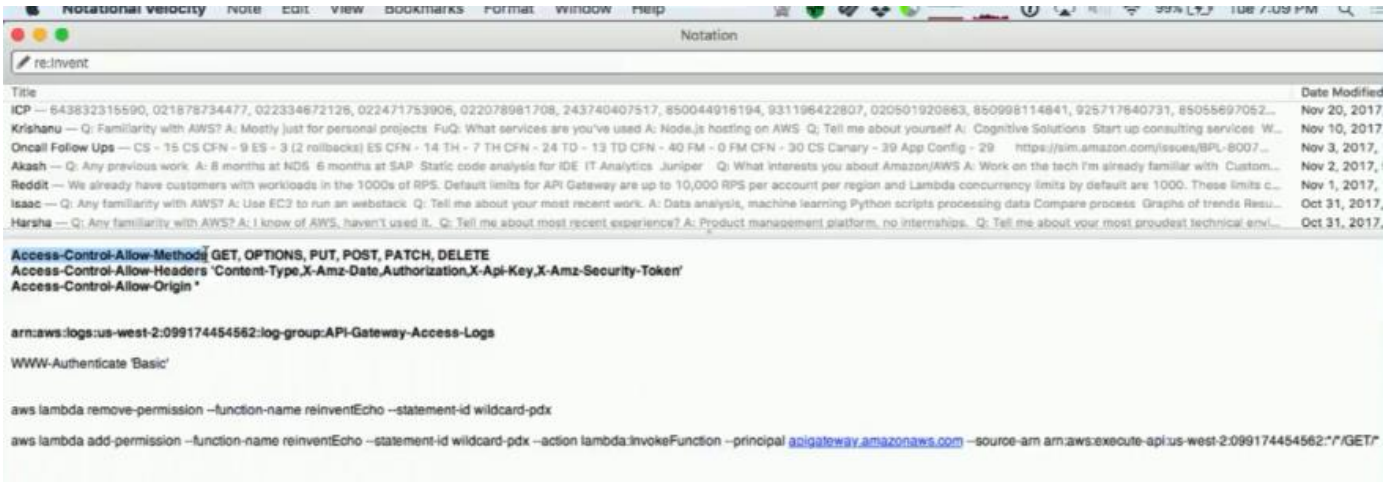


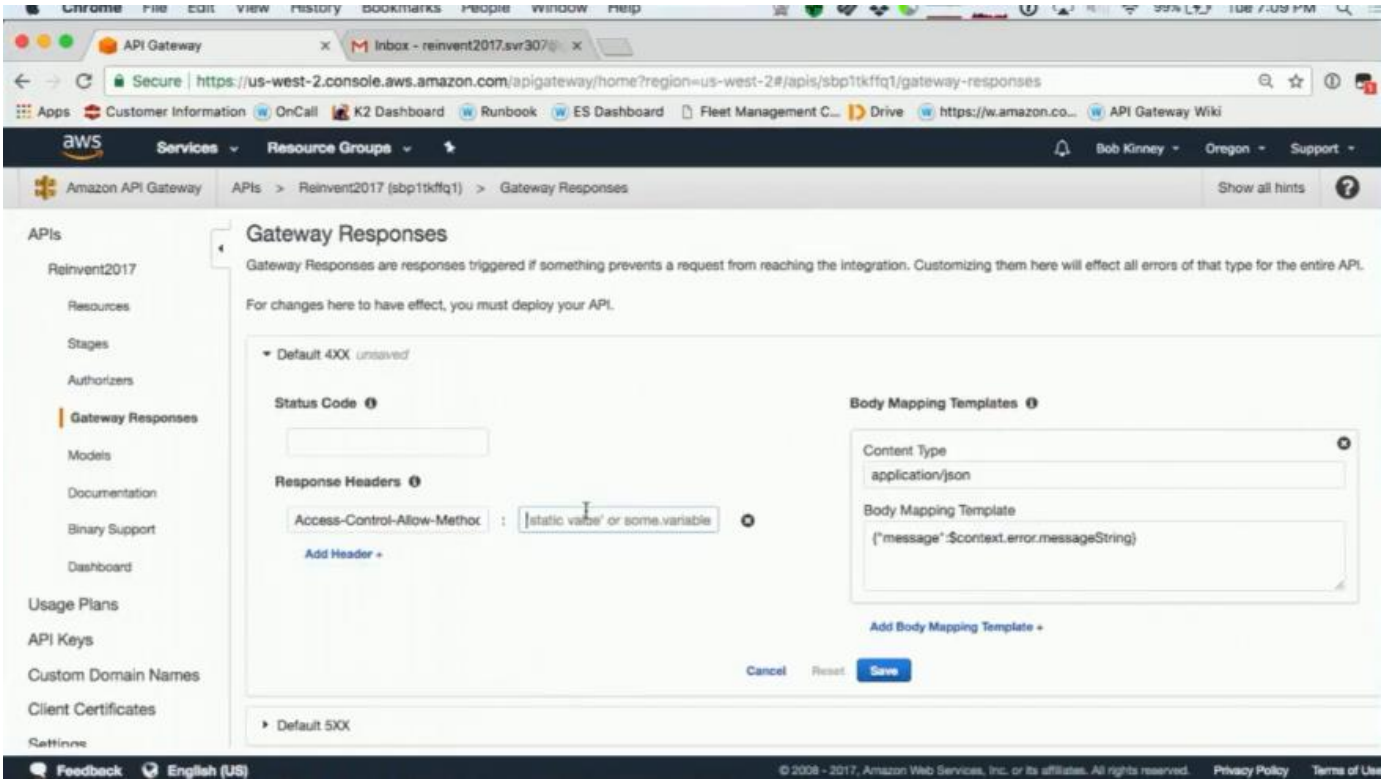

API Gateway
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs
Show all hints
APIs
+ Create API
Reinvent2017
No description.
Endpoint Configuration
Endpoint Type
Edge Optimized
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Settings
Feedback
English (US)
Privacy Policy
Terms of Use


API Gateway
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/resources/kzxft12k52
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Resources > / (kzxft12k52)
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Resources
Actions
/
GET
/auth
/custom
/allow
GET
/deny
GET
/gatling
/random
GET
/throttle
GET
/unauth
GET
/ Methods
GET
arn:aws:lambda:us-west-2:099174454562:functi...
Authorization None
API Key Not required
Feedback
English (US)
Privacy Policy
Terms of Use

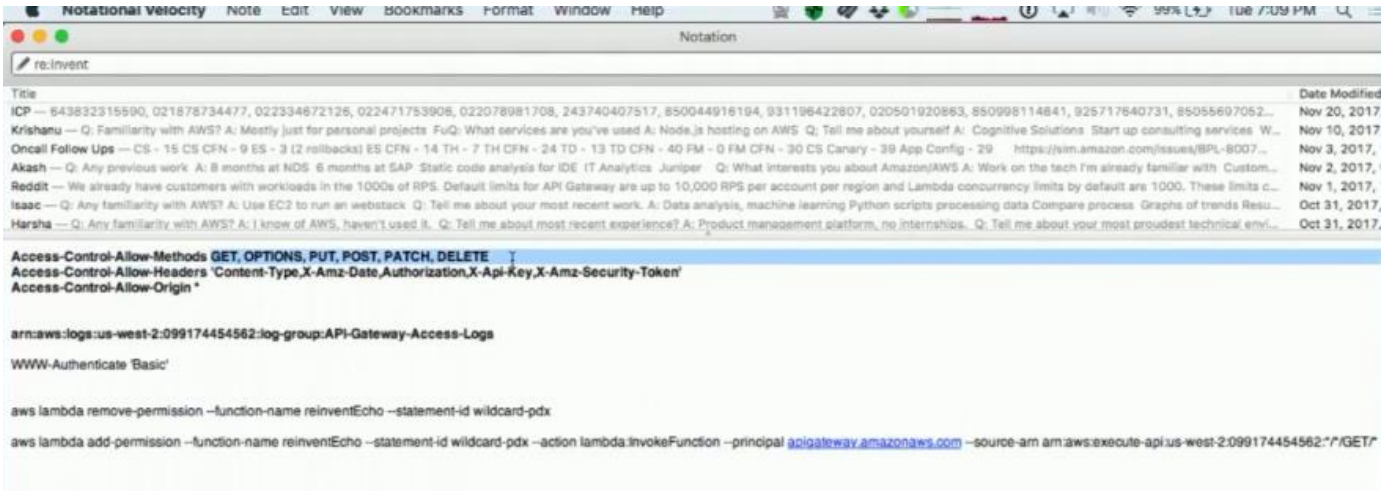


We simply want to add CORS headers to all our error messages and so we need to override all our default 4xx and 5xx responses to do this.

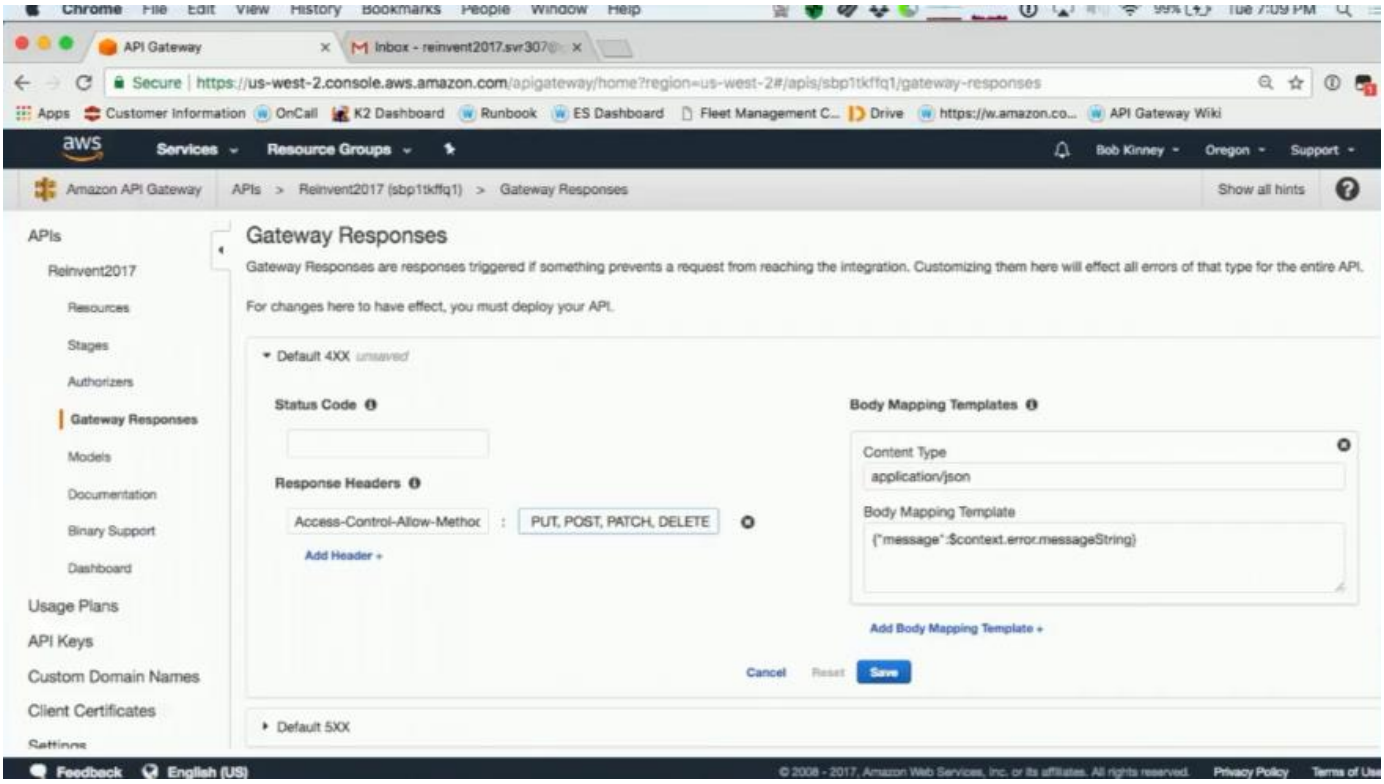


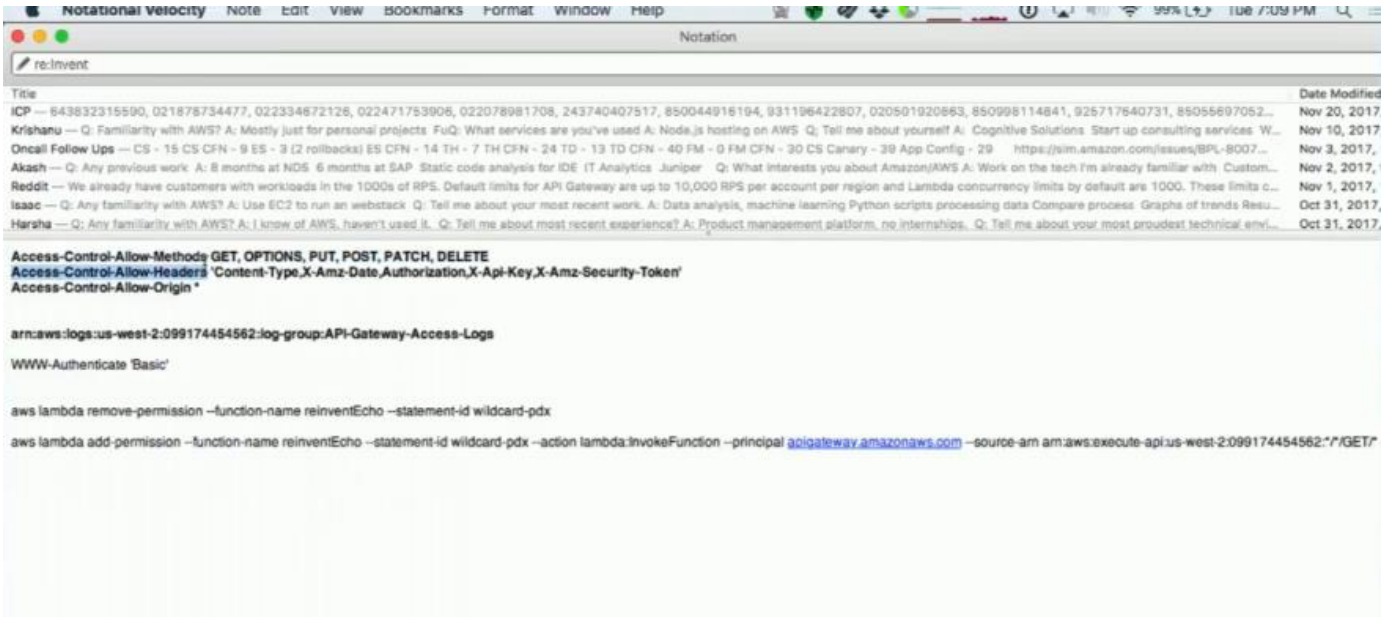


We add our headers as above for CORS

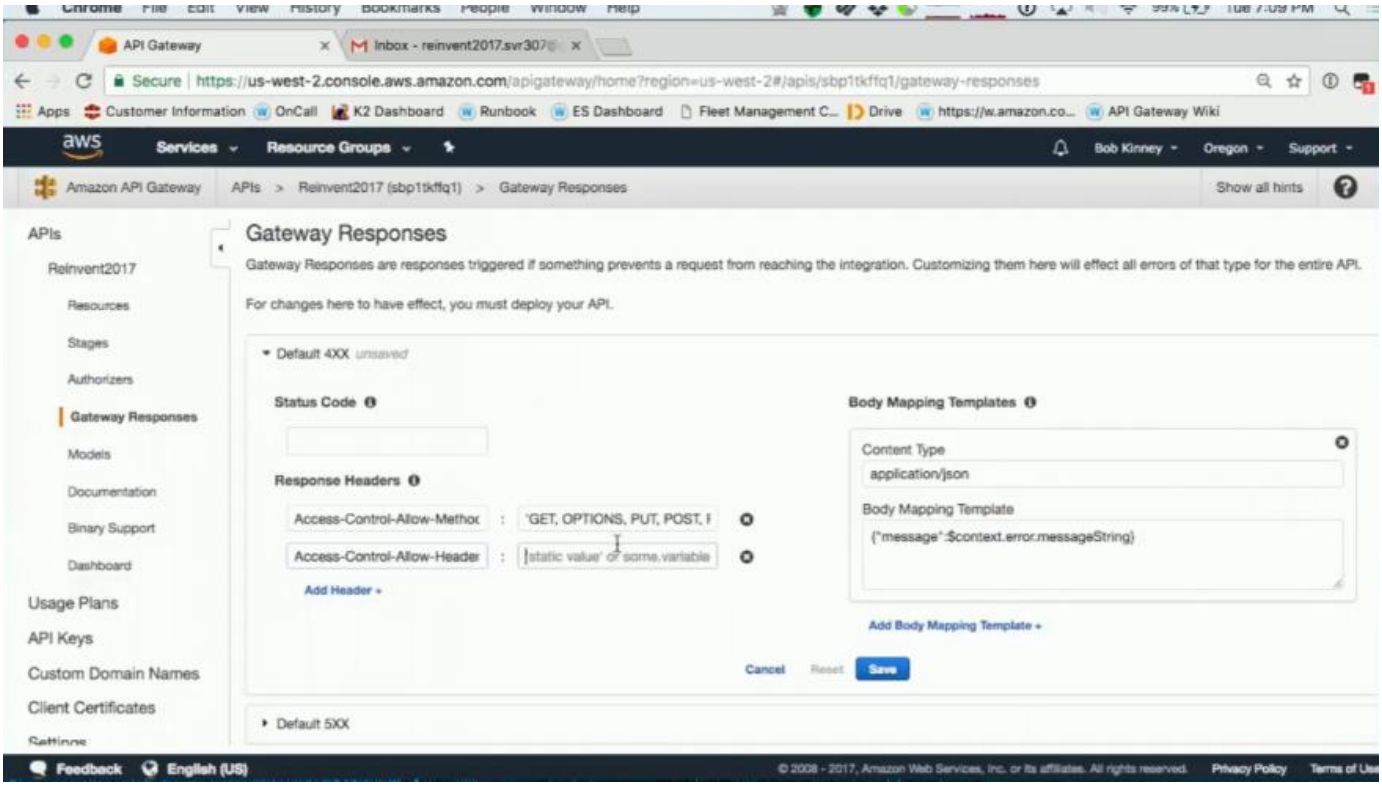


Then we add the messages too

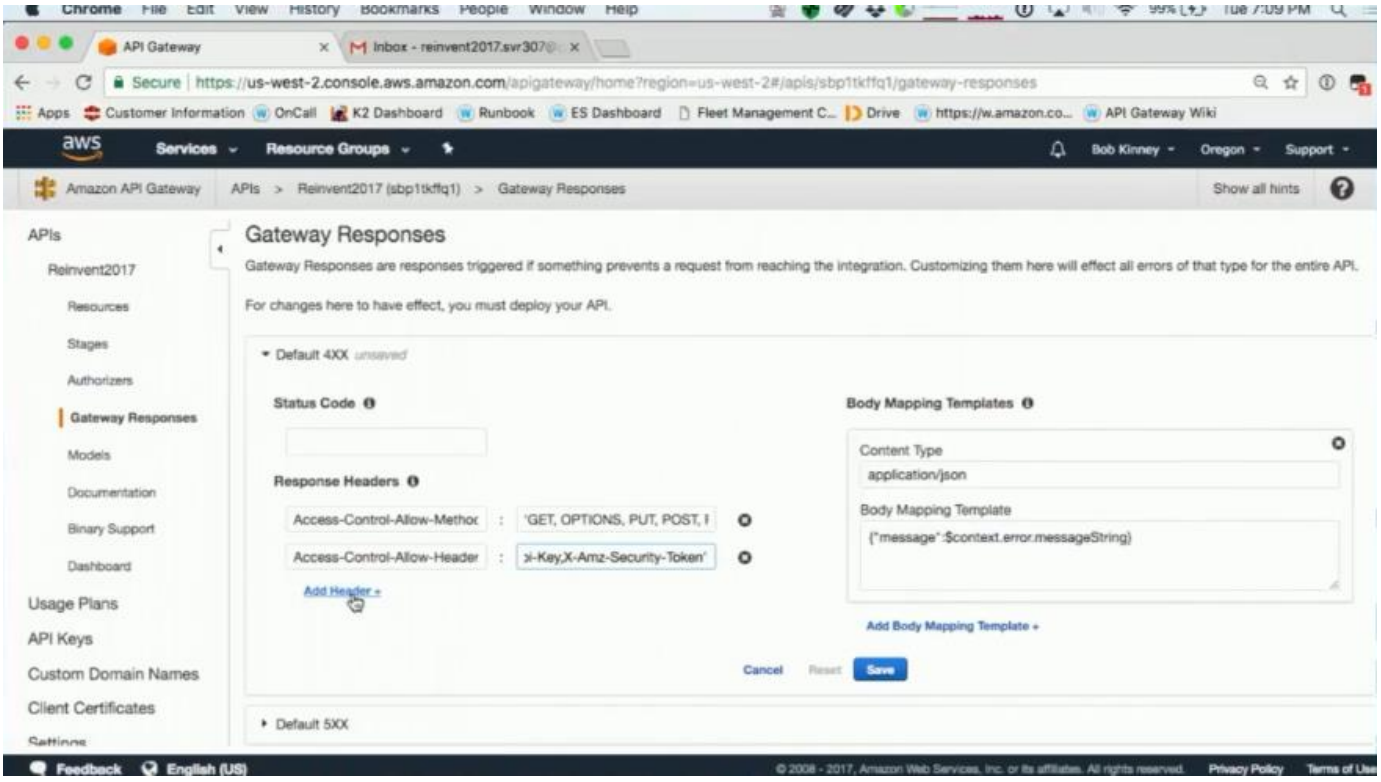




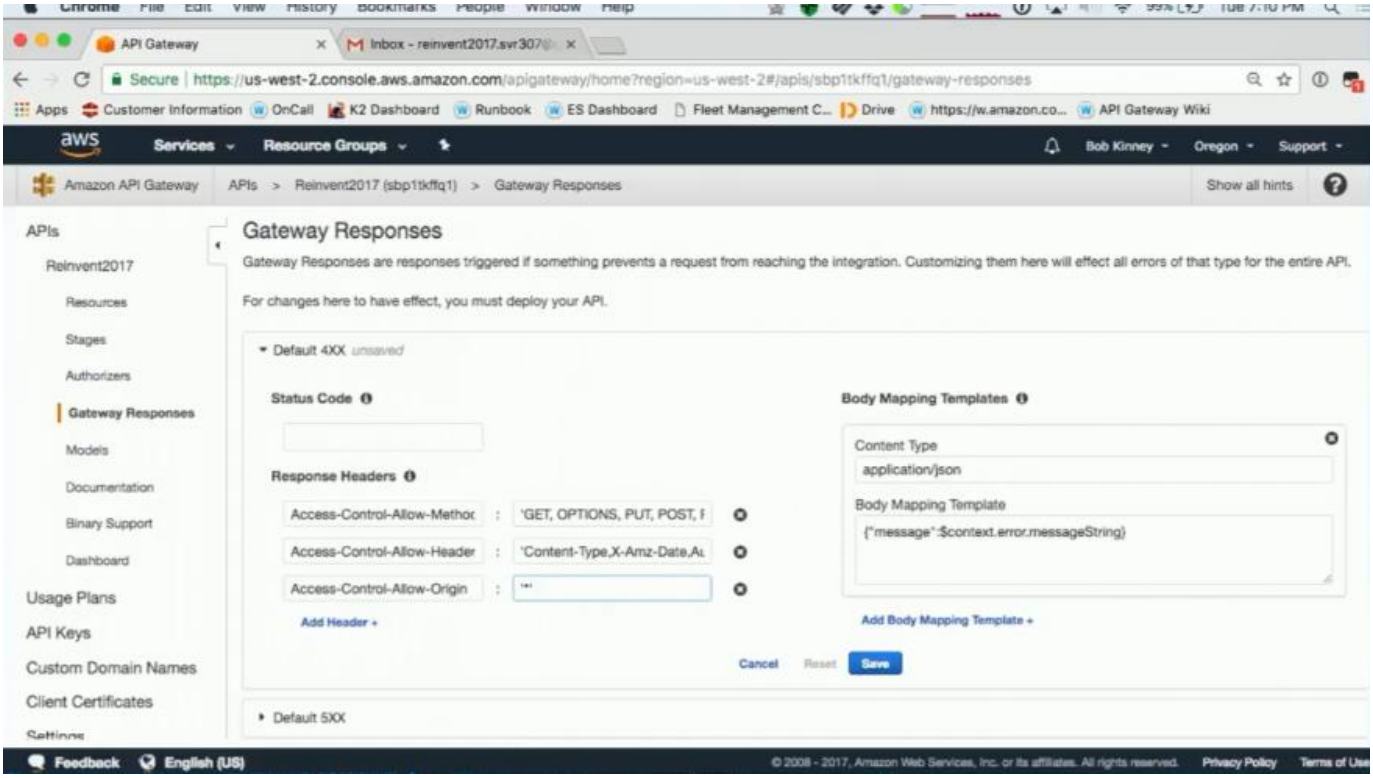
We then add our access control headers too


API Gateway
Inbox - reinvent2017.svr307
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/gateway-responses
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
Amazon API Gateway APIs > Reinvent2017 (sbp1tkffq1) > Gateway Responses Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Settings
Gateway Responses
Gateway Responses are responses triggered if something prevents a request from reaching the integration. Customizing them here will effect all errors of that type for the entire API.
For changes here to have effect, you must deploy your API.
Default 4XX unsaved
Status Code
Response Headers
Access-Control-Allow-Methoc : 'GET, OPTIONS, PUT, POST, F
Access-Control-Allow-Header : 'static value' or some.variable
Add Header +
Body Mapping Templates
Content Type
application/json
Body Mapping Template
{"message":$context.error.messageString}
Add Body Mapping Template +
Cancel Reset Save
Default 5XX
Feedback English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

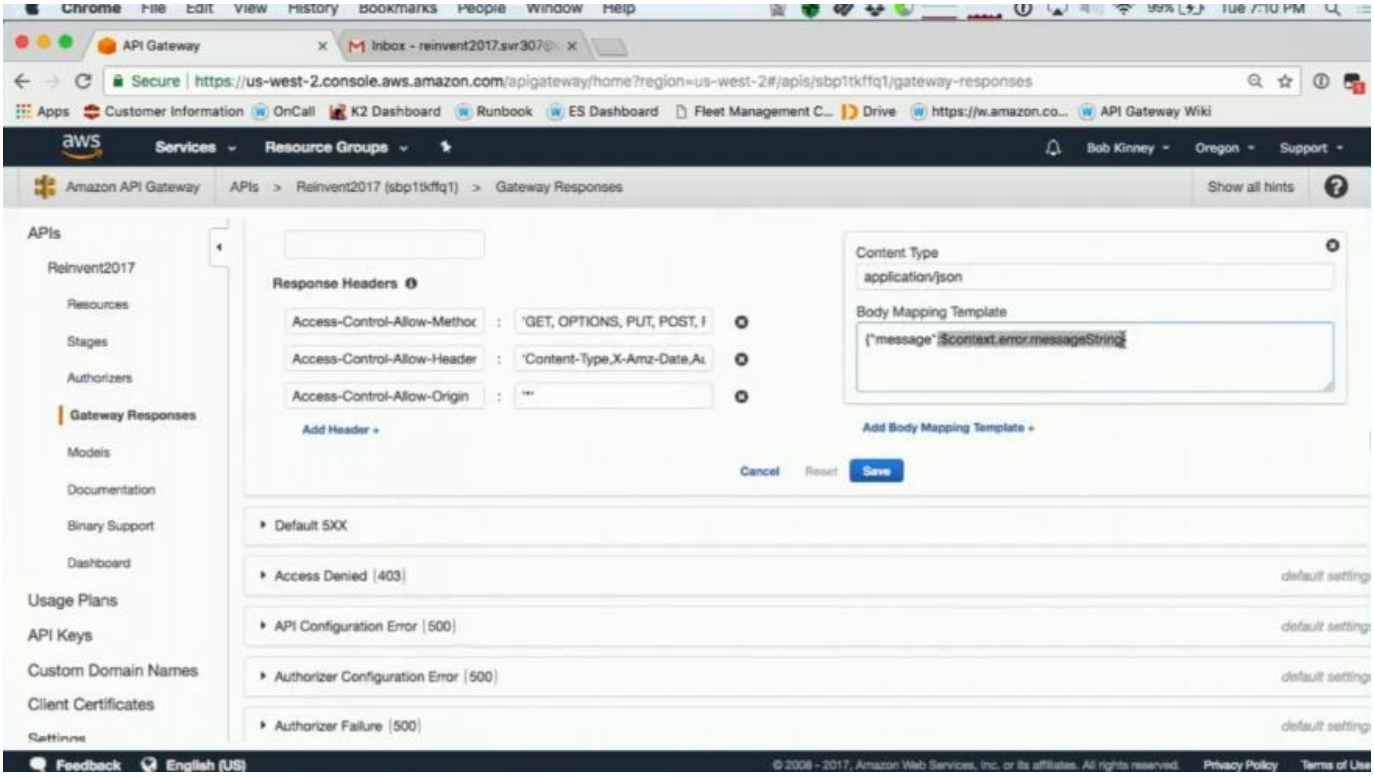

Notational Velocity Note Edit View Bookmarks Format Window Help
Notation
re:Invent
Title
Date Modified
ICP — 643832315590, 021878734477, 022334672126, 022471753906, 022078981708, 243740407517, 850044916194, 931196422807, 020501920863, 850998114841, 925717640731, 850556970... Nov 20, 2017,
Krishanu — Q: Familiarity with AWS? A: Mostly just for personal projects FuQ: What services are you've used A: Node.js hosting on AWS Q: Tell me about yourself A: Cognitive Solutions Start up consulting services W... Nov 10, 2017,
Oncall Follow Ups — CS - 15 CS CFN - 9 ES - 3 (2 rollbacks) ES CFN - 14 TH - 7 TH CFN - 24 TD - 13 TD CFN - 40 FM - 0 FM CFN - 30 CS Canary - 39 App Config - 29 https://sim.amazon.com/issues/BPL-8007... Nov 3, 2017,
Akash — Q: Any previous work A: 8 months at NDS 6 months at SAP Static code analysis for IDE IT Analytics Juniper Q: What interests you about Amazon/AWS A: Work on the tech I'm already familiar with Custom... Nov 2, 2017,
Reddit — We already have customers with workloads in the 1000s of RPS. Default limits for API Gateway are up to 10,000 RPS per account per region and Lambda concurrency limits by default are 1000. These limits c... Nov 1, 2017,
Isaac — Q: Any familiarity with AWS? A: Use EC2 to run an webstack Q: Tell me about your most recent work. A: Data analysis, machine learning Python scripts processing data Compare process Graphs of trends Resu... Oct 31, 2017,
Harsha — Q: Any familiarity with AWS? A: I know of AWS, haven't used it. Q: Tell me about most recent experience? A: Product management platform, no internships. Q: Tell me about your most proudest technical envi... Oct 31, 2017,
Access-Control-Allow-Methods GET, OPTIONS, PUT, POST, PATCH, DELETE
Access-Control-Allow-Headers 'Content-Type,X-Amz-Date,Authorization,X-Api-Key,X-Amz-Security-Token'
Access-Control-Allow-Origin *
arn:aws:logs:us-west-2:099174454562:log-group:API-Gateway-Access-Logs
WWW-Authenticate 'Basic'
aws lambda remove-permission --function-name reinventEcho --statement-id wildcard-pdx
aws lambda add-permission --function-name reinventEcho --statement-id wildcard-pdx --action lambda:InvokeFunction --principal apigateway.amazonaws.com --source-arn arn:aws:execute-api:us-west-2:099174454562:*/*/GET/*


API Gateway
Inbox - reinvent2017.svr307@
https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/gateway-responses
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Gateway Responses
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Gateway Responses
Gateway Responses are responses triggered if something prevents a request from reaching the integration. Customizing them here will effect all errors of that type for the entire API.
For changes here to have effect, you must deploy your API.
Default 4XX unsaved
Status Code
Response Headers
Access-Control-Allow-Methoc : 'GET, OPTIONS, PUT, POST, I
Access-Control-Allow-Header : xi-Key,X-Amz-Security-Token'
Add Header +
Body Mapping Templates
Content Type
application/json
Body Mapping Template
{"message":$context.error.messageString}
Add Body Mapping Template +
Cancel
Reset
Save
Default 5XX
Feedback
English (US)
Privacy Policy
Terms of Use

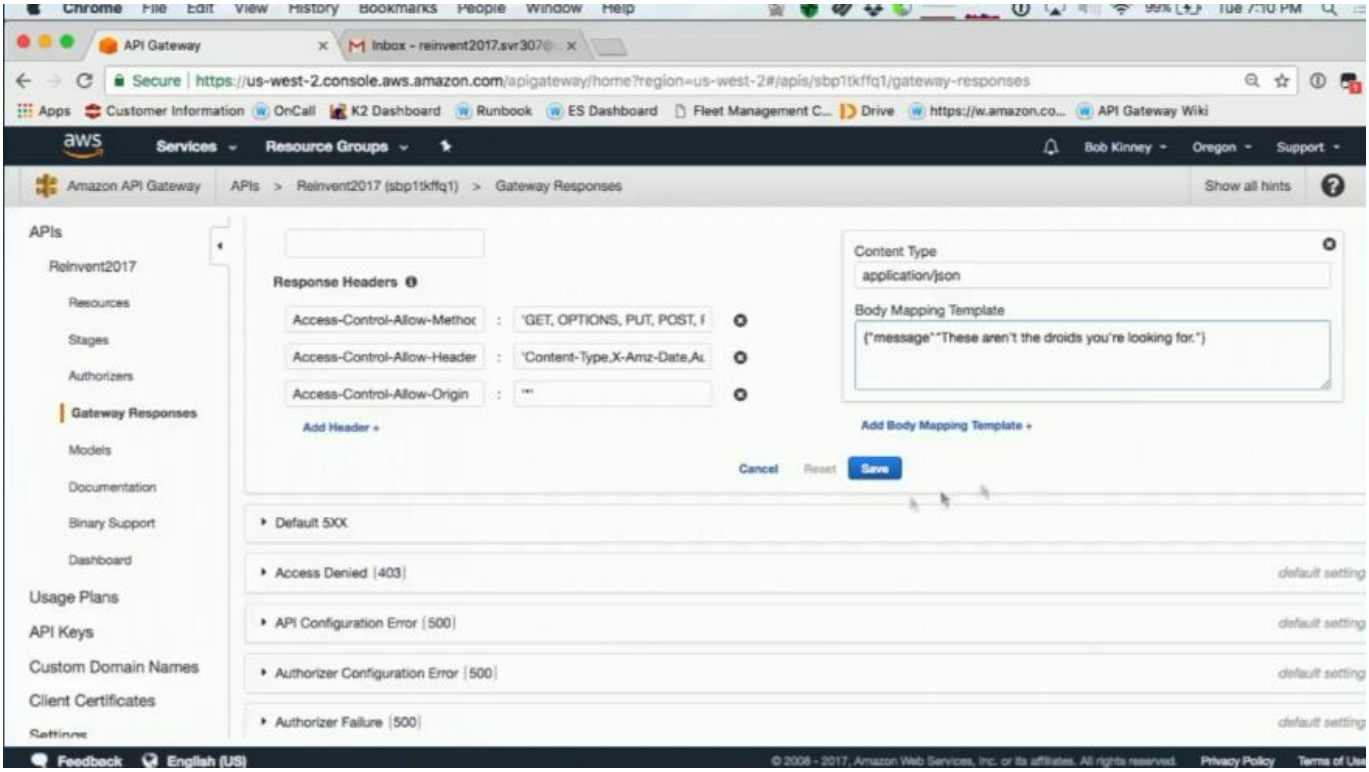

Notation
re:invent
Title
Date Modified
Access-Control-Allow-Methods GET, OPTIONS, PUT, POST, PATCH, DELETE
Access-Control-Allow-Headers 'Content-Type,X-Amz-Date,Authorization,X-Api-Key,X-Amz-Security-Token'
Access-Control-Allow-Origin *
arn:aws:logs:us-west-2:099174454562:log-group:API-Gateway-Access-Logs
WWW-Authenticate 'Basic'
aws lambda remove-permission --function-name reinventEcho --statement-id wildcard-pdx
aws lambda add-permission --function-name reinventEcho --statement-id wildcard-pdx --action lambda:InvokeFunction --principal apigateway.amazonaws.com --source-arn arn:aws:execute-api:us-west-2:099174454562:*/*/GET/*



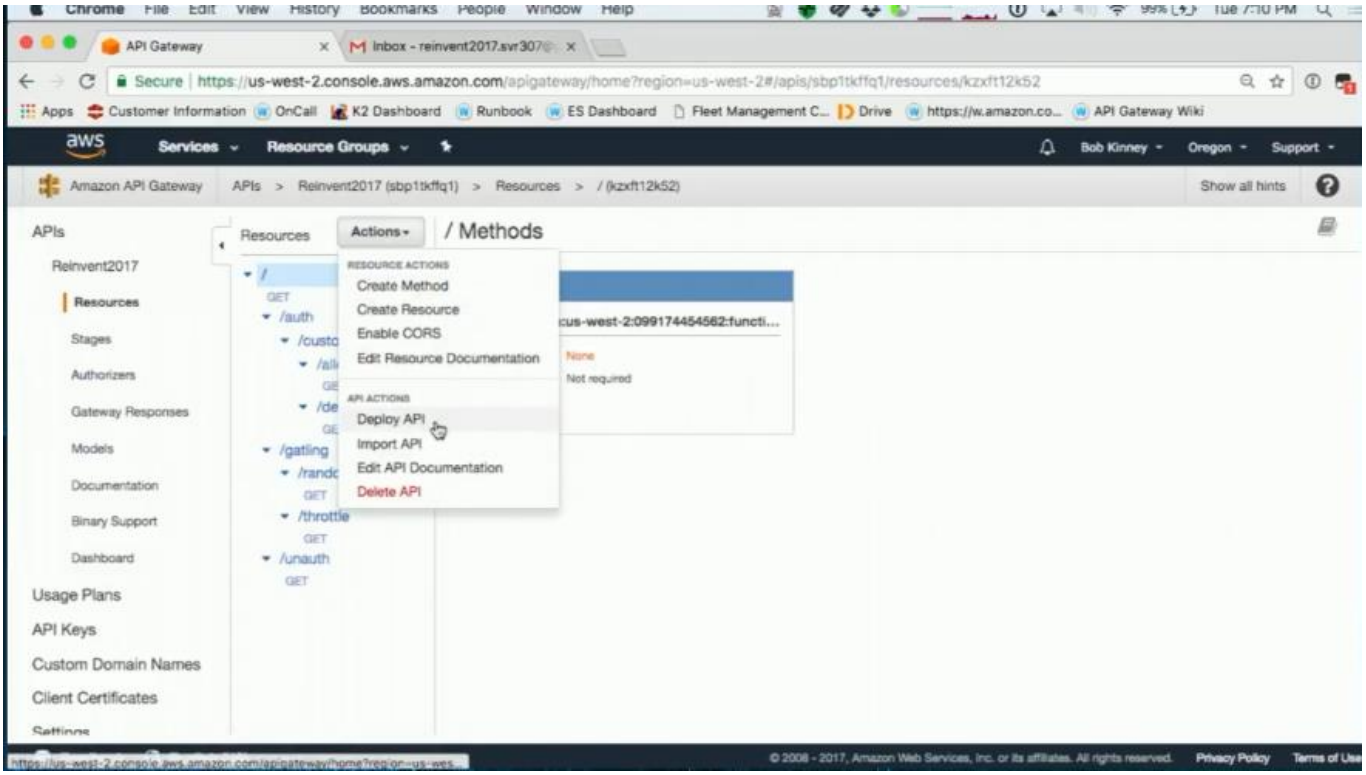
We then add our origin. We can also lock this down to a specific domain like www.ssceexams.com instead of using the generic * as above

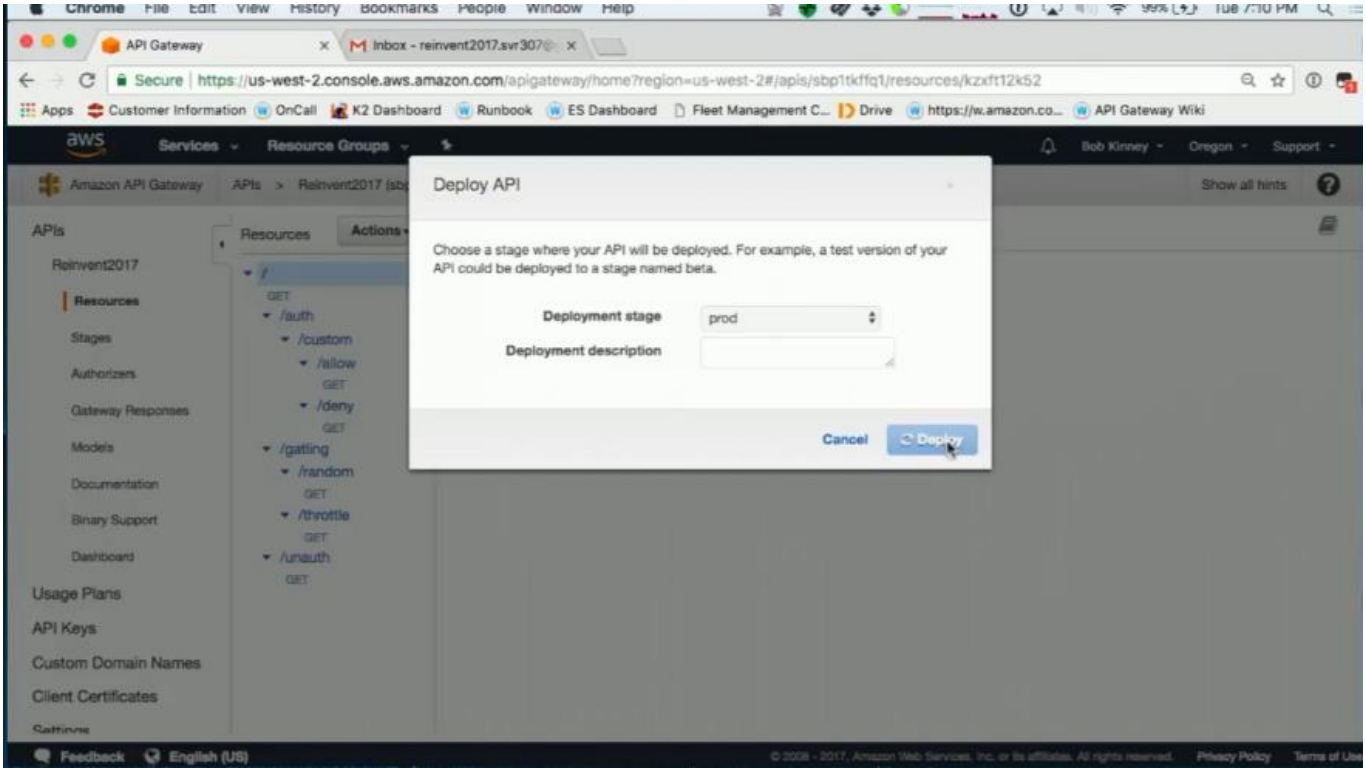


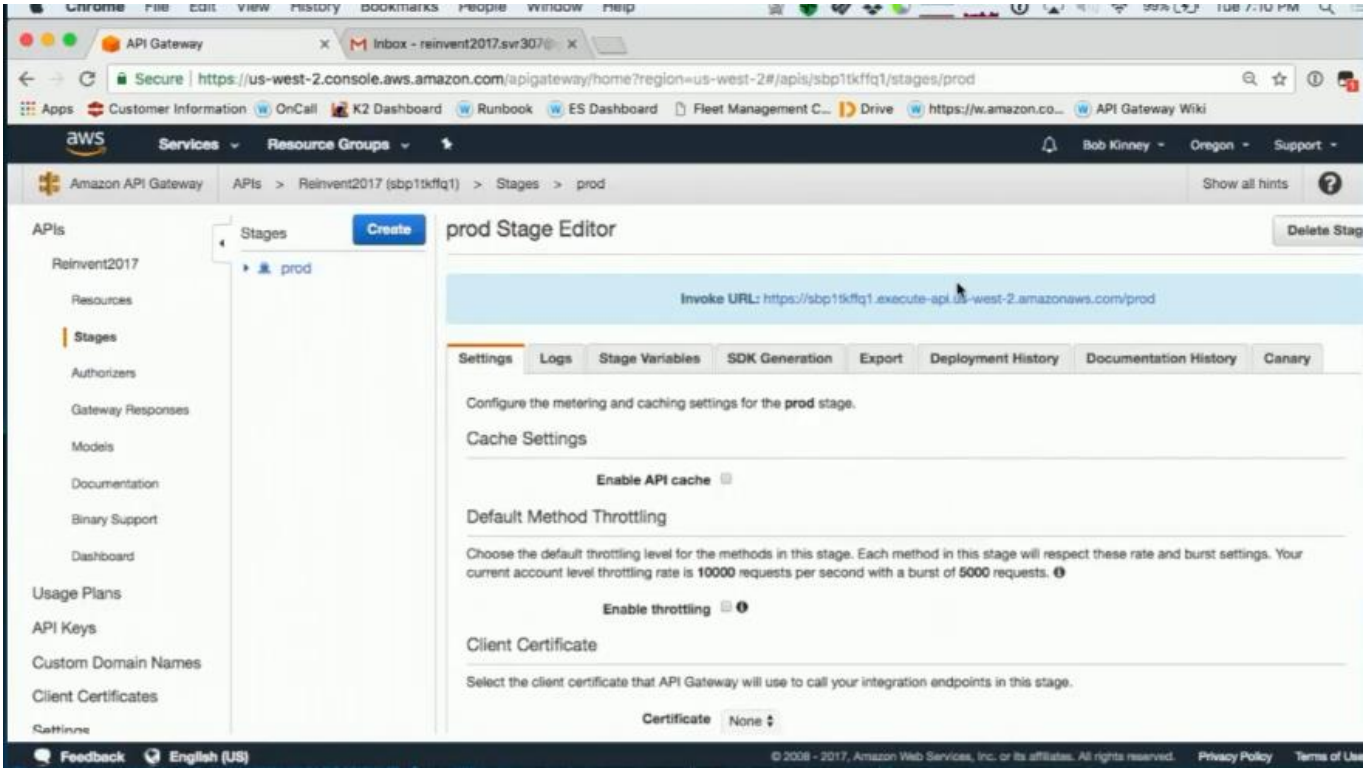
Now we can additionally go ahead and update the error message to be displayed as above

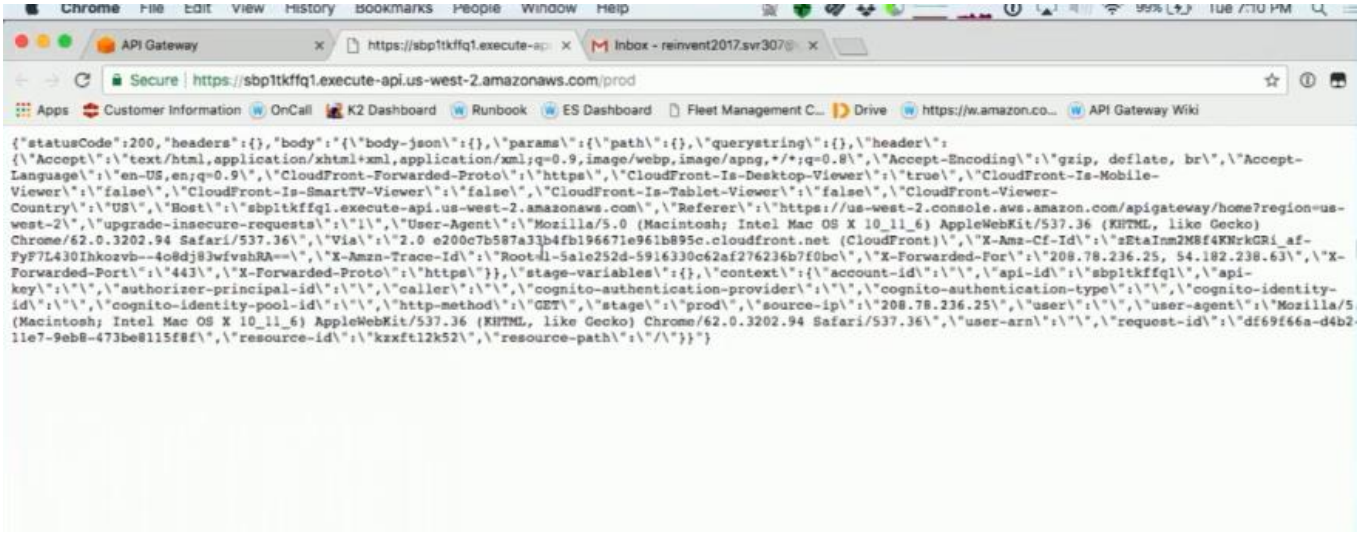


We then click the Save button to update our API Gateway response for 4xx and 5xx errors. This will save it into the API Gateway service, we still need to deploy these changes as below

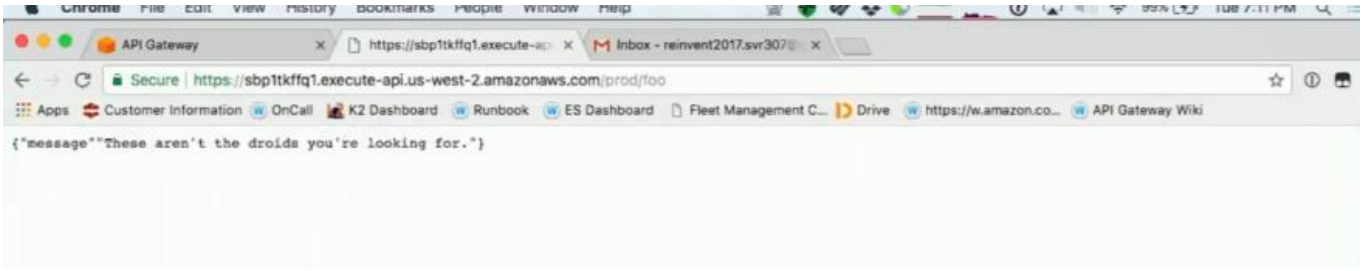



Deploy API
Choose a stage where your API will be deployed. For example, a test version of your API could be deployed to a stage named beta.
Deployment stage
prod
Deployment description
Cancel
Deploy
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates

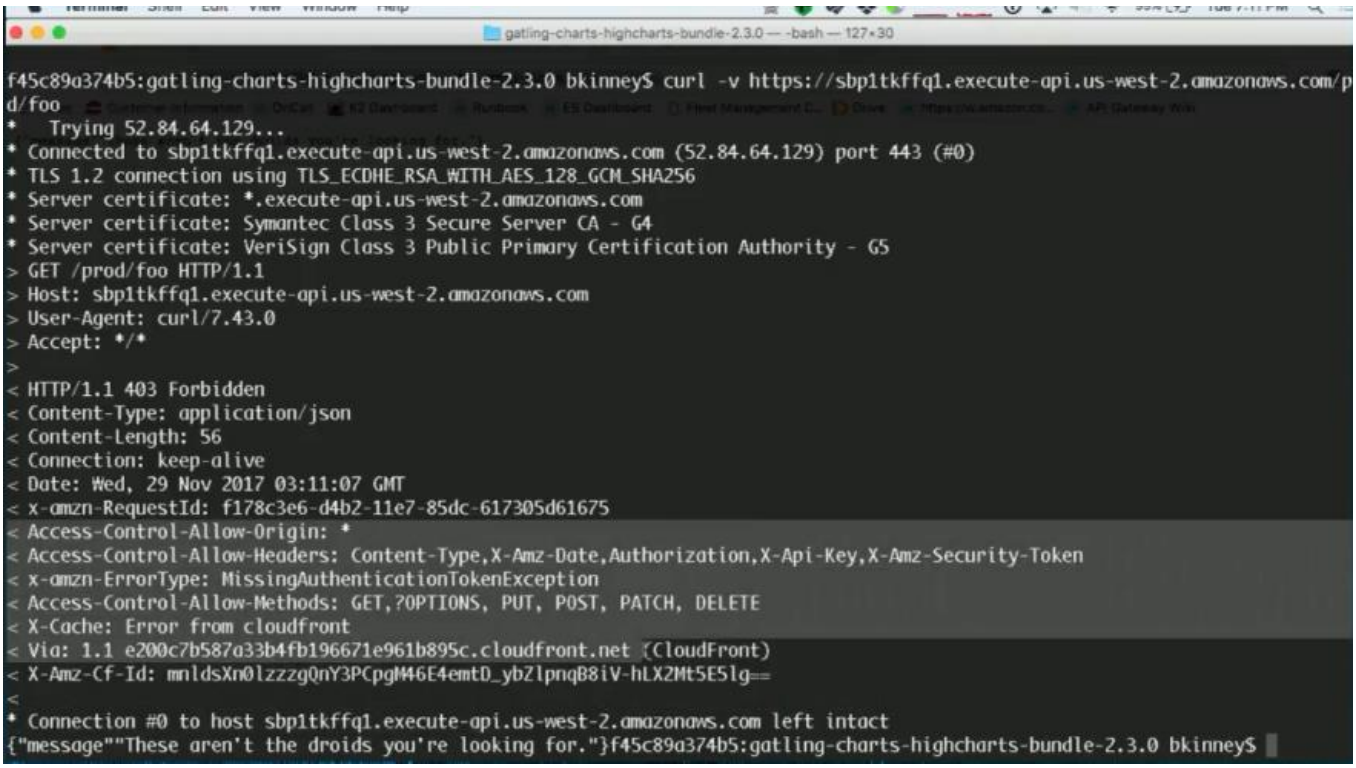

prod Stage Editor
Invoke URL: https://sbp1tkffq1.execute-api.us-west-2.amazonaws.com/prod
Settings
Logs
Stage Variables
SDK Generation
Export
Deployment History
Documentation History
Canary
Configure the metering and caching settings for the prod stage.
Cache Settings
Enable API cache
Default Method Throttling
Choose the default throttling level for the methods in this stage. Each method in this stage will respect these rate and burst settings. Your current account level throttling rate is 10000 requests per second with a burst of 5000 requests.
Enable throttling
Client Certificate
Select the client certificate that API Gateway will use to call your integration endpoints in this stage.
Certificate
None



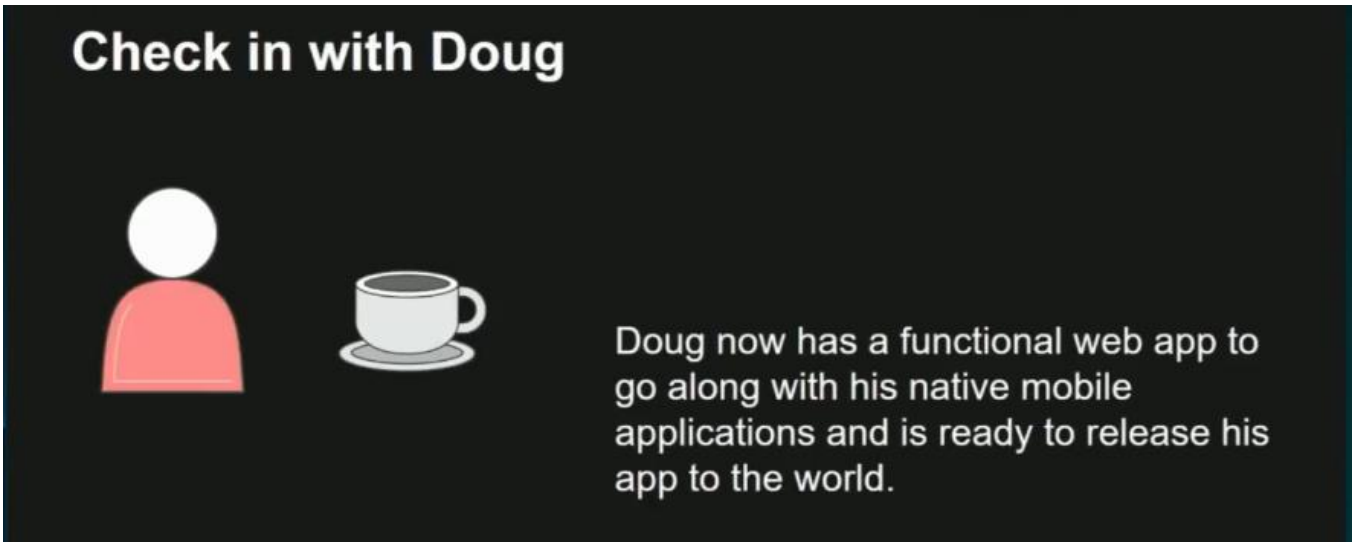
Using the API Gateway endpoint, we can see what a 200 response looks like with some actual values as above.



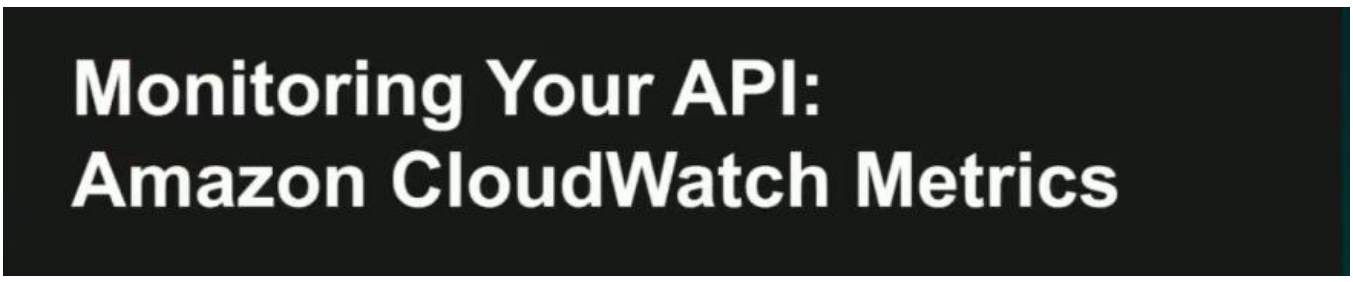
We can use a path that is not mapped in our API Gateway to test the error message for 4xx and 5xx as above. We can see that we are getting back our custom error message



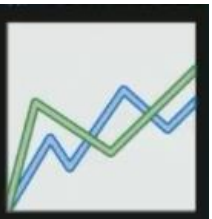
Using curl, we can also see that we are also getting our CORS headers that we set earlier. This shows us how to customize the error responses for all the error cases for our API Gateway.



Now that Doug is able to get CORS working with custom error handling, he is able to build his JS web app and ready to go live



Amazon CloudWatch Metrics



API Gateway Default metrics set

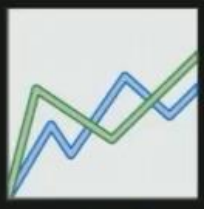
- Count—total number of invokes received by API Gateway
- 4XXError—number of invokes that generated a 4XX error (includes throttling)
- 5XXError—number of invokes that generated a 5XX error
- Latency—total time API Gateway took to fully process request
- IntegrationLatency—time API Gateway took to call integration
- CacheHitCount—number of successful cache fetches
- CacheMissCount—number of unsuccessful cache fetches

Amazon CloudWatch Metrics

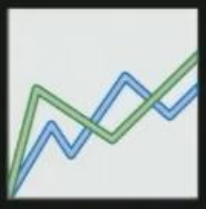
Detailed metrics

- Same set of metrics at method level
- Can be enabled globally or only for specific methods

GET



PUT



DELETE

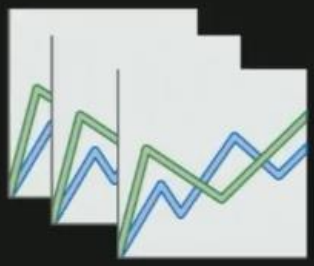
Amazon CloudWatch Metrics

Default Metrics

Included for free
Broken down by API stage

Detailed Metrics

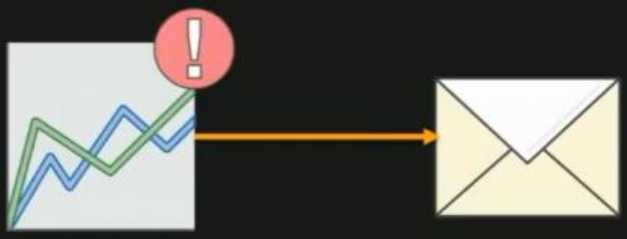
Standard CloudWatch pricing
Broken down by method



Amazon CloudWatch Alarms

Any metric can be tied to an alarm
Alarm notifications can be sent to Amazon SNS topic
SNS topic can then send to any number of destinations

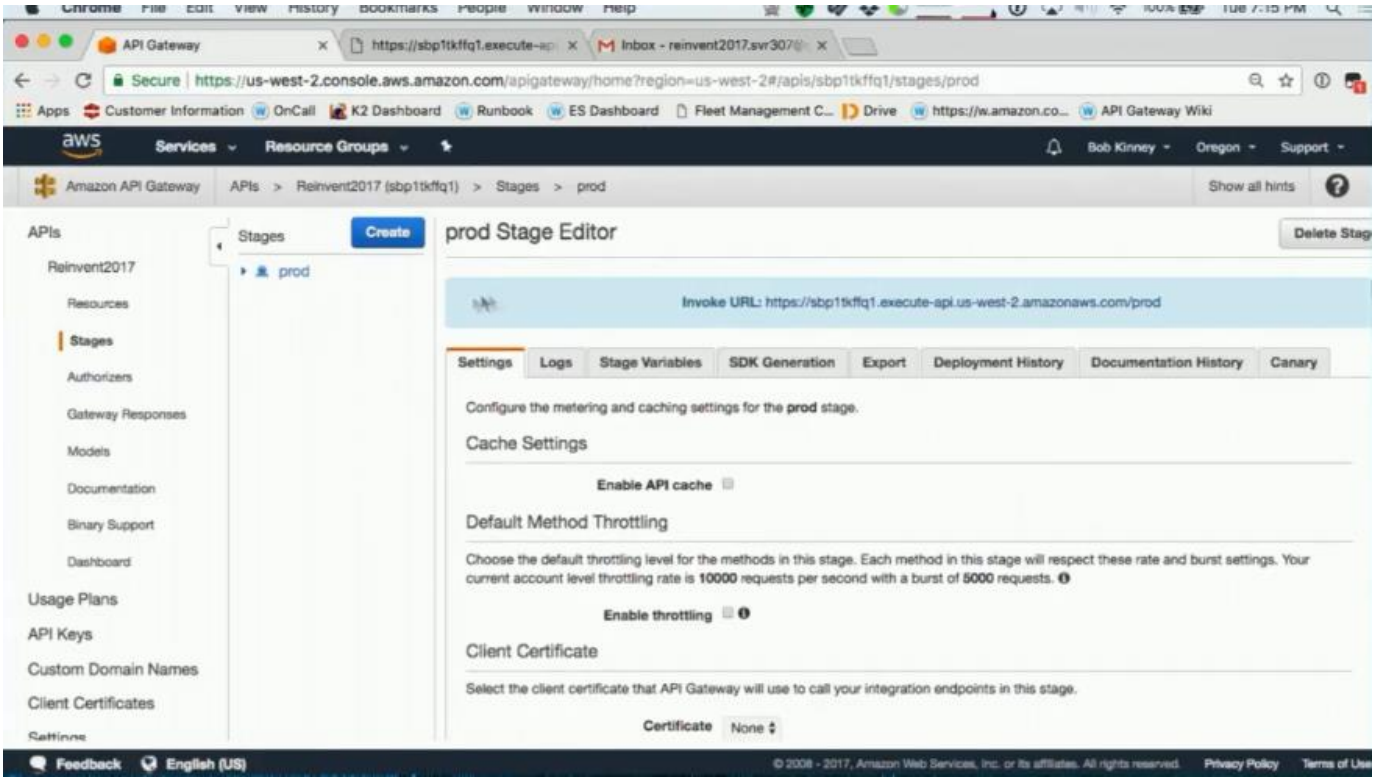
- E-mail address
- SQS queue
- Lambda Function



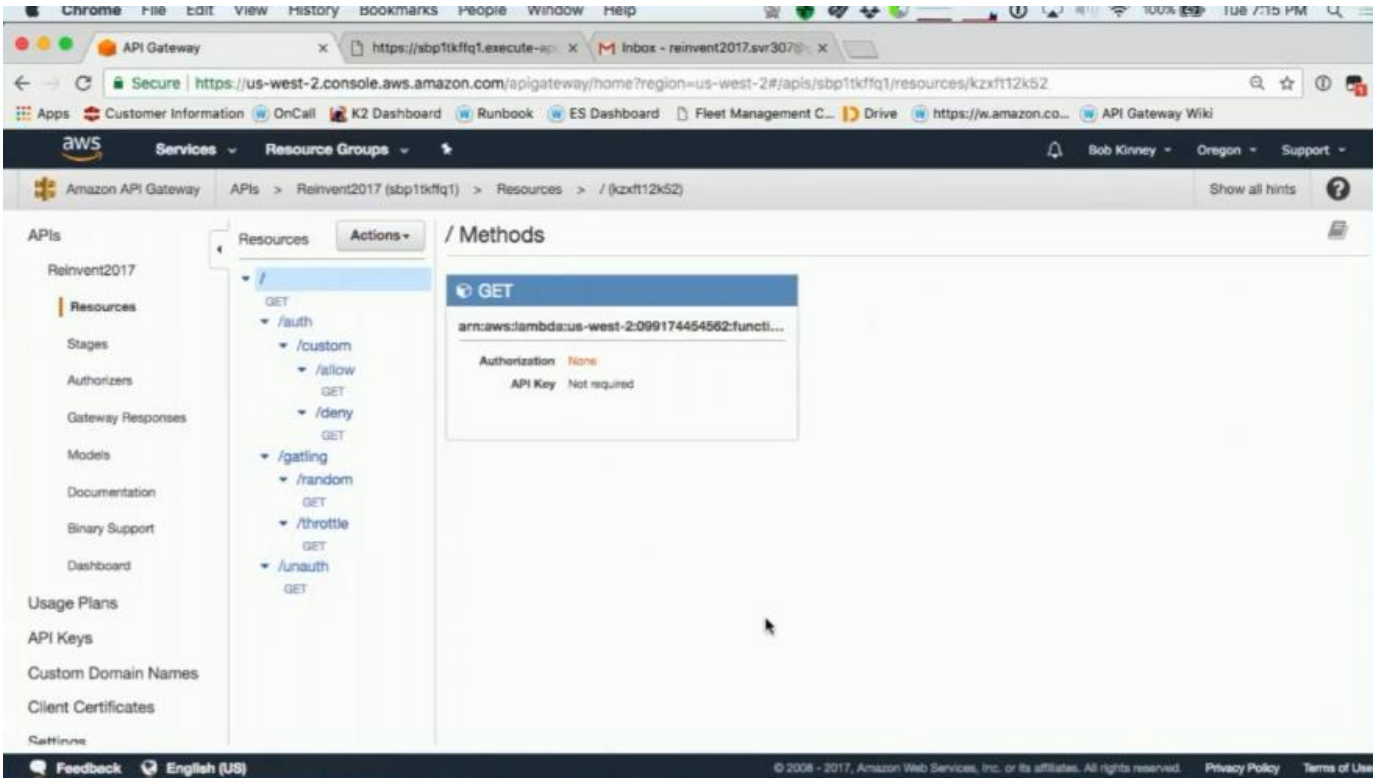
Once we have metrics, then we are able to set alarms on those metrics.

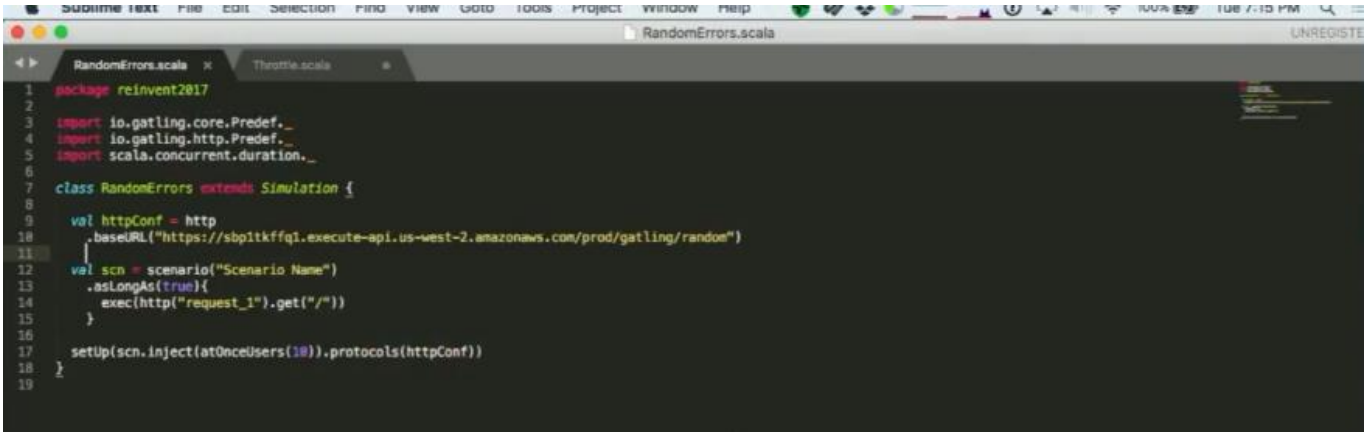
Demo

Let us now see how we can set up some alarms based on our CloudWatch metrics.

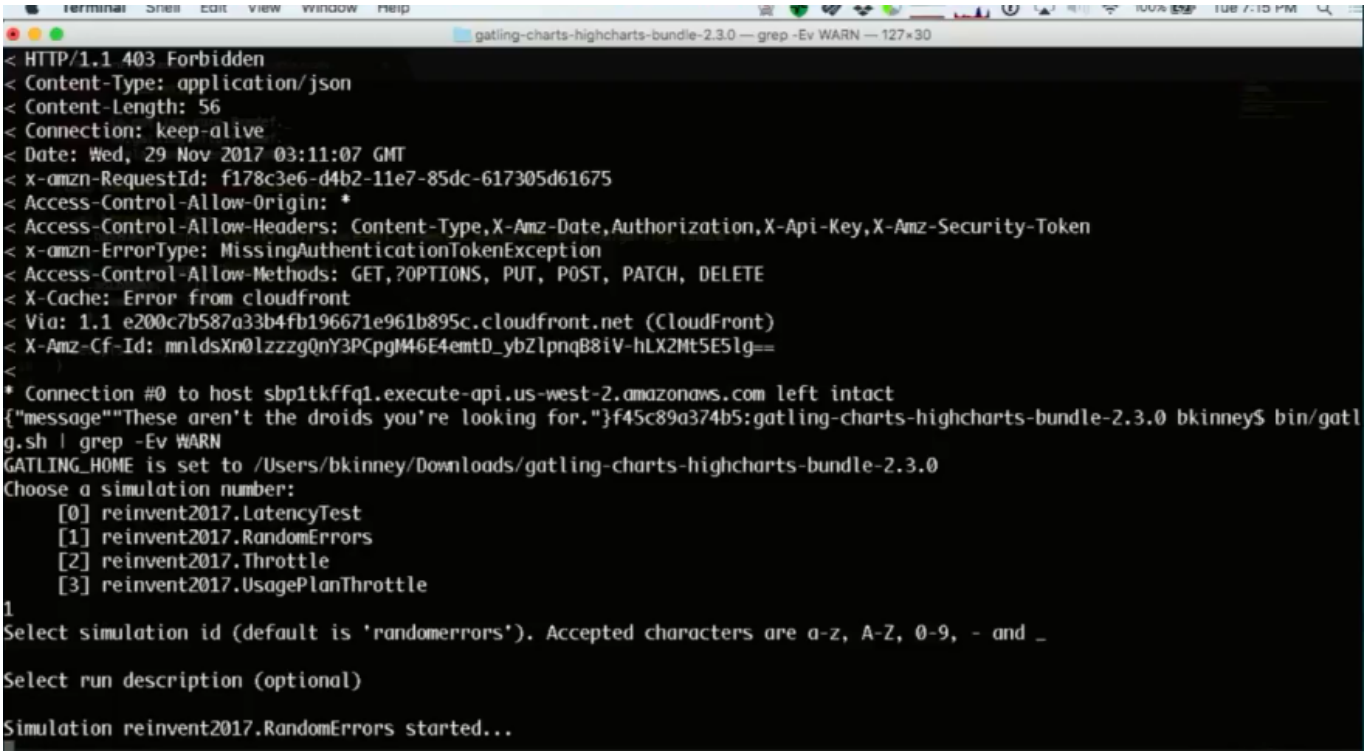


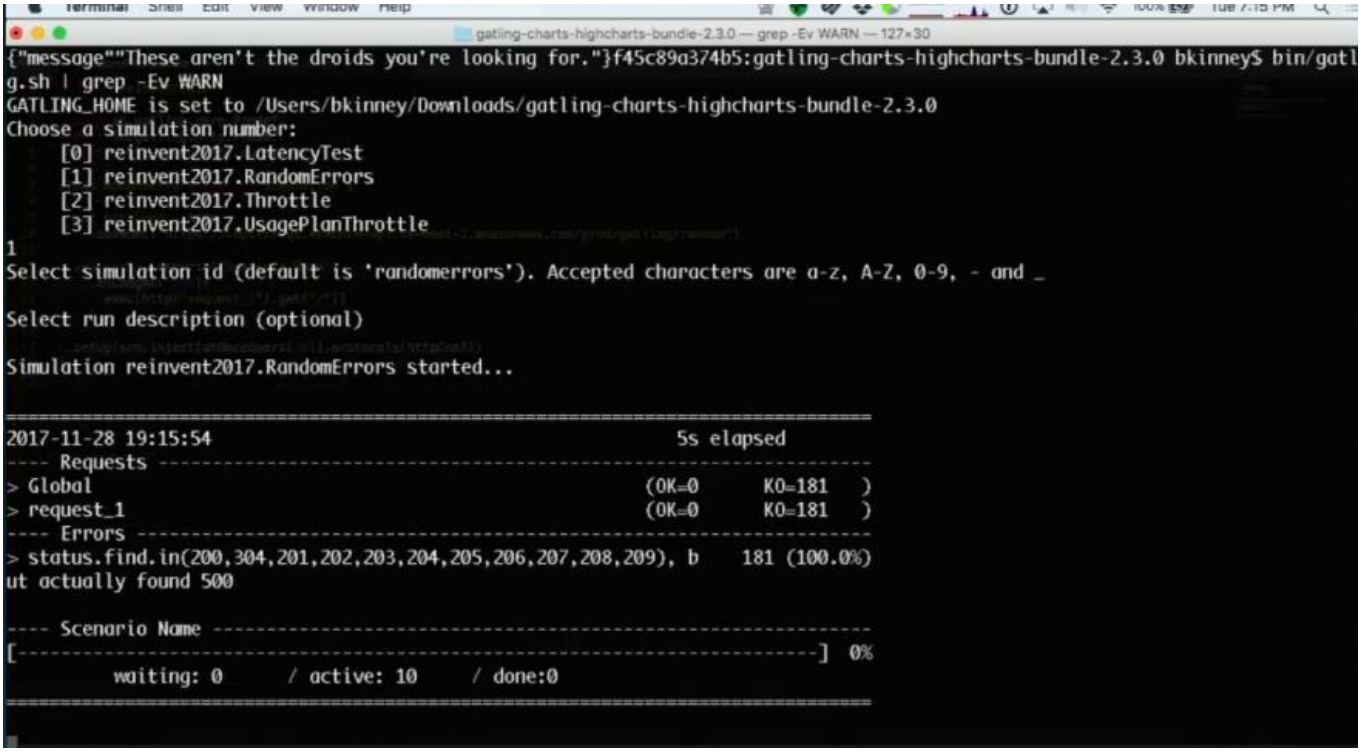
We still have our API Gateway as above

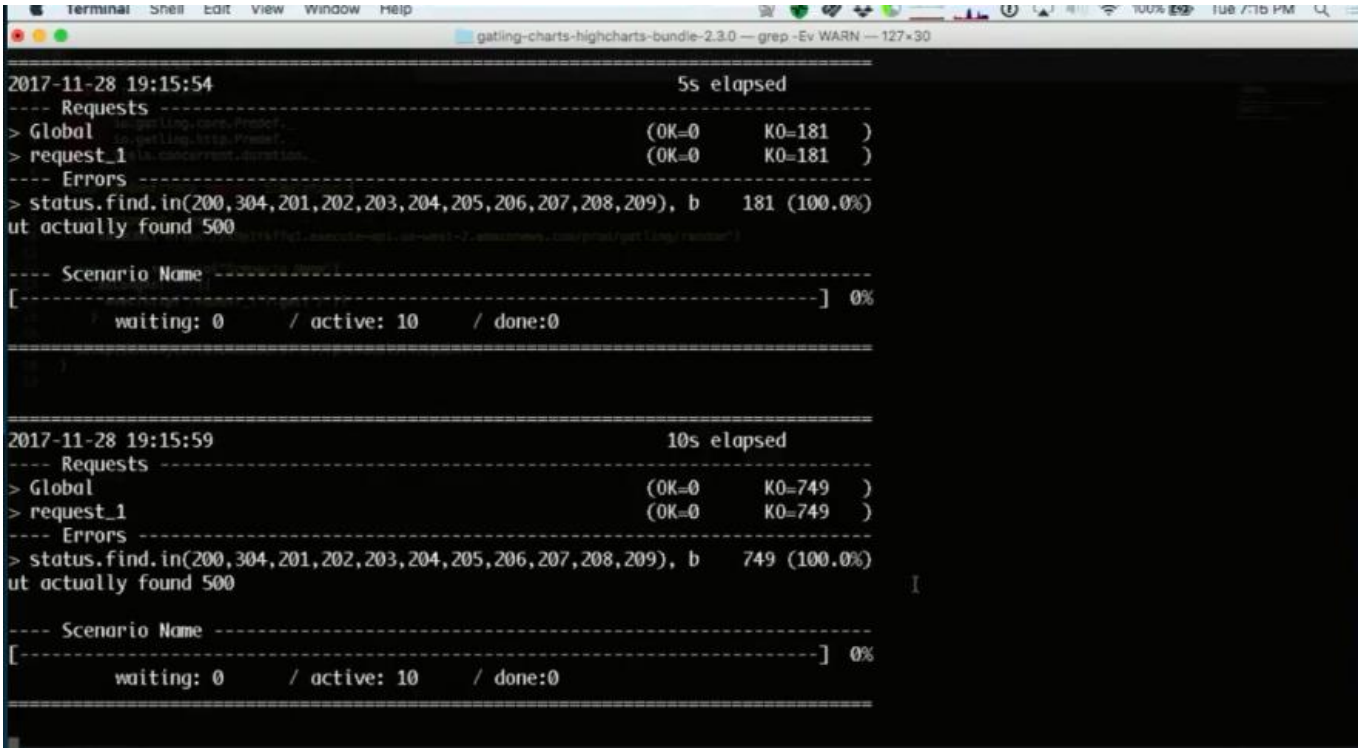


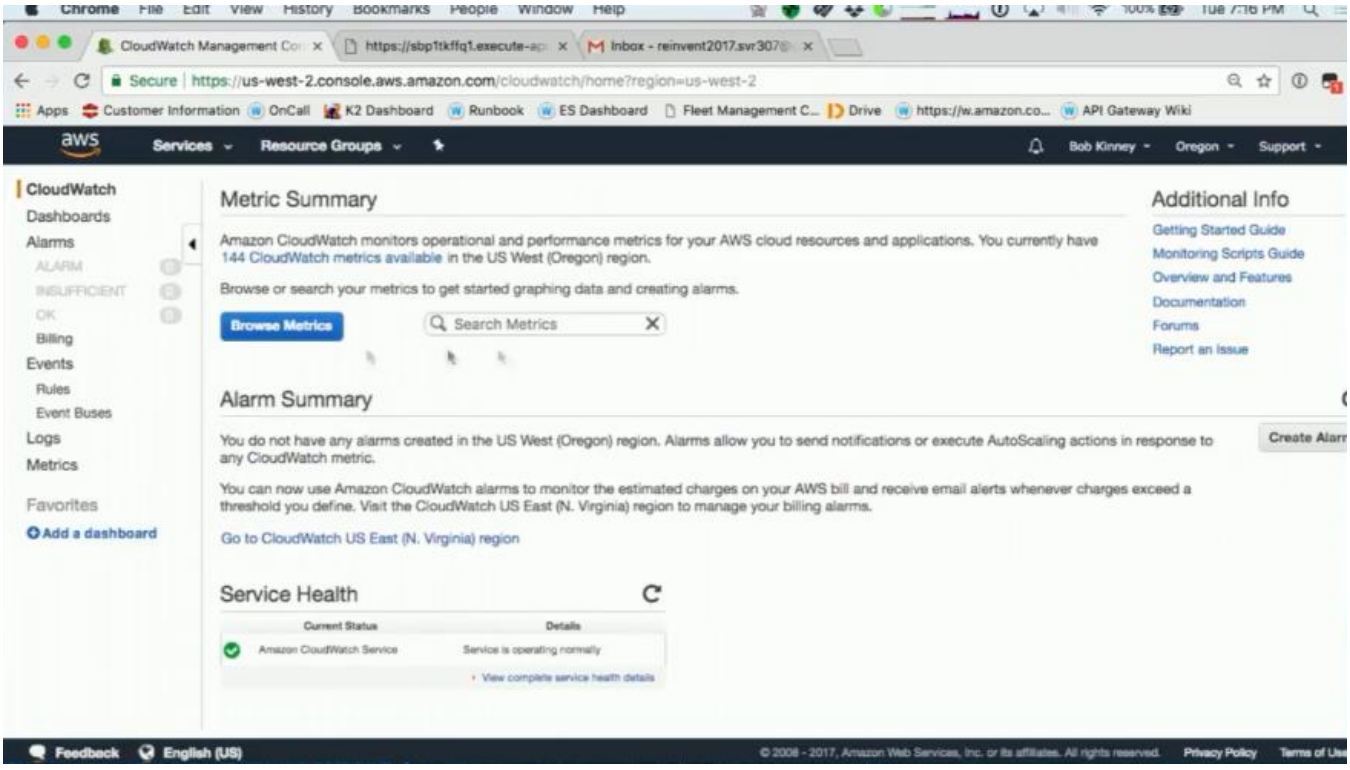


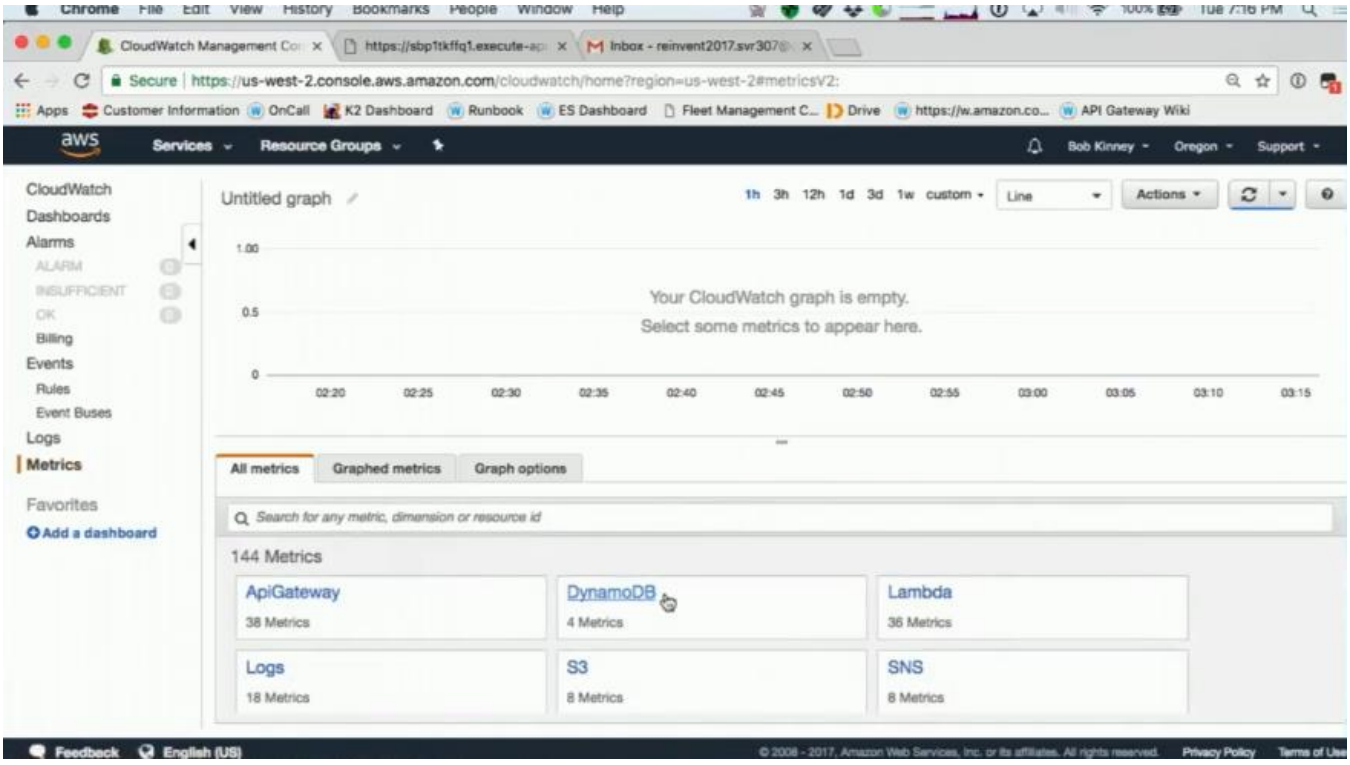
We can use a tool called Gatlin to generate some traffic to our API. This will hit a particular lambda function that will generate different error messages

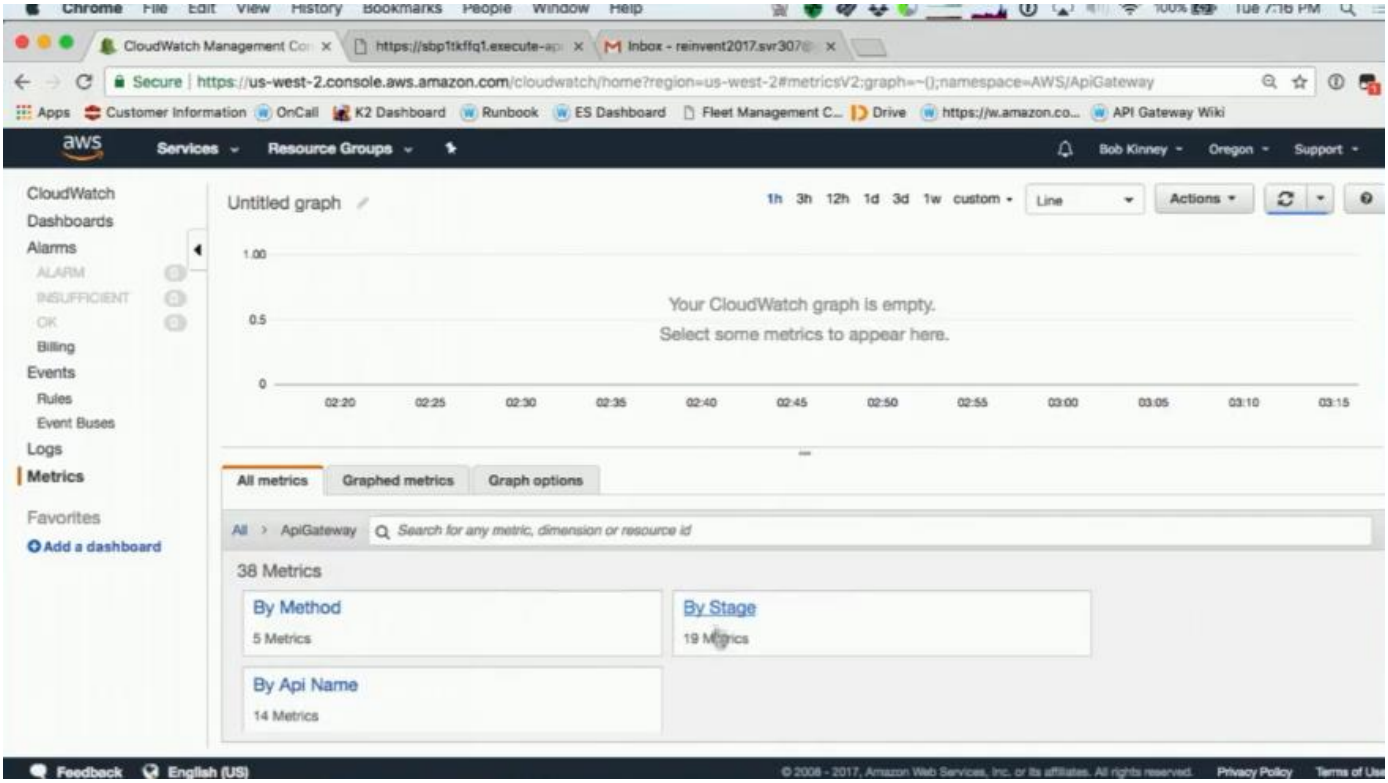


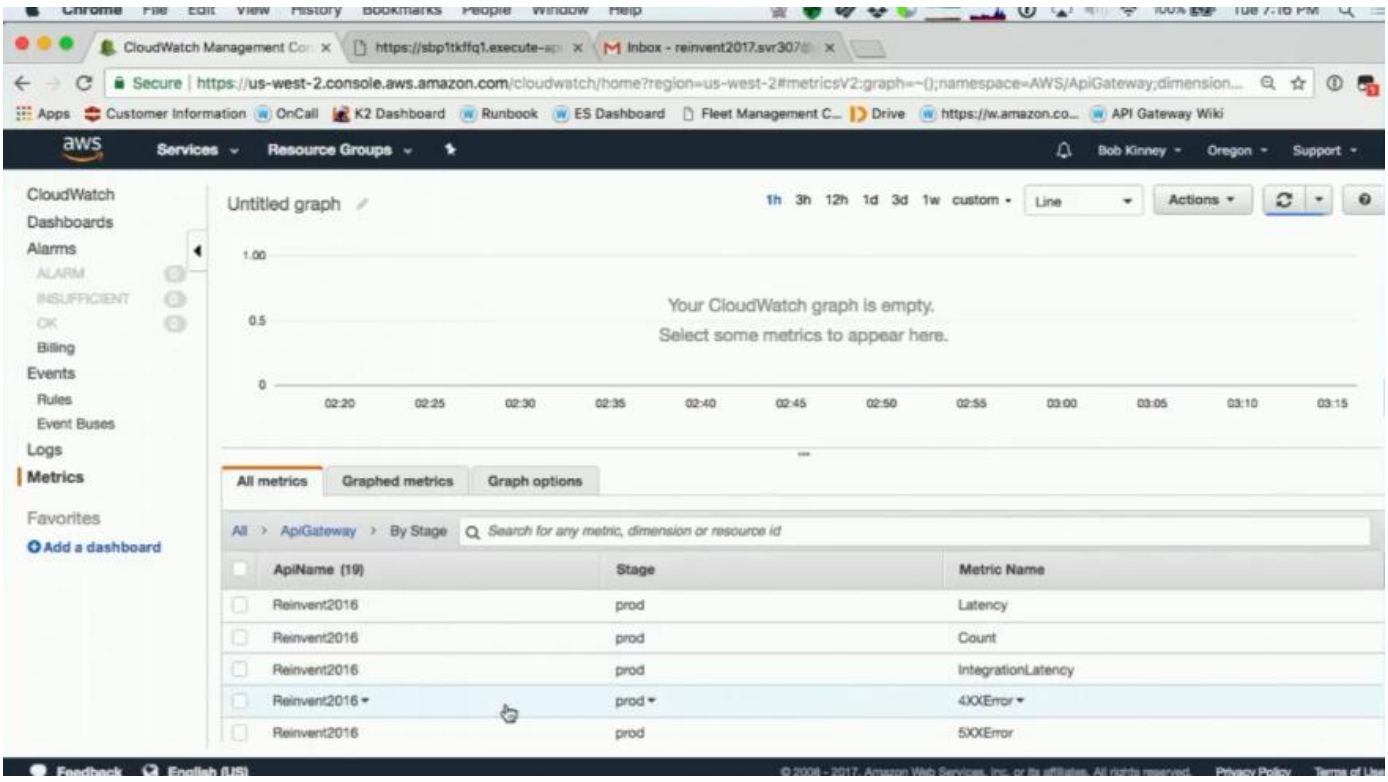

gatling-charts-highcharts-bundle-2.3.0 — grep -Ev WARN — 127×30
{"message""These aren't the droids you're looking for."}f45c89a374b5:gatling-charts-highcharts-bundle-2.3.0 bkinney$ bin/gatl
g.sh | grep -Ev WARN
GATLING_HOME is set to /Users/bkinney/Downloads/gatling-charts-highcharts-bundle-2.3.0
Choose a simulation number:
[0] reinvent2017.LatencyTest
[1] reinvent2017.RandomErrors
[2] reinvent2017.Throttle
[3] reinvent2017.UsagePlanThrottle
1
Select simulation id (default is 'randomerrors'). Accepted characters are a-z, A-Z, 0-9, - and _
Select run description (optional)
Simulation reinvent2017.RandomErrors started...
2017-11-28 19:15:54 5s elapsed
---- Requests
> Global (OK=0 KO=181)
> request_1 (OK=0 KO=181)
---- Errors
> status.find.in(200,304,201,202,203,204,205,206,207,208,209), b 181 (100.0%)
ut actually found 500
---- Scenario Name
[] 0%
waiting: 0 / active: 10 / done:0

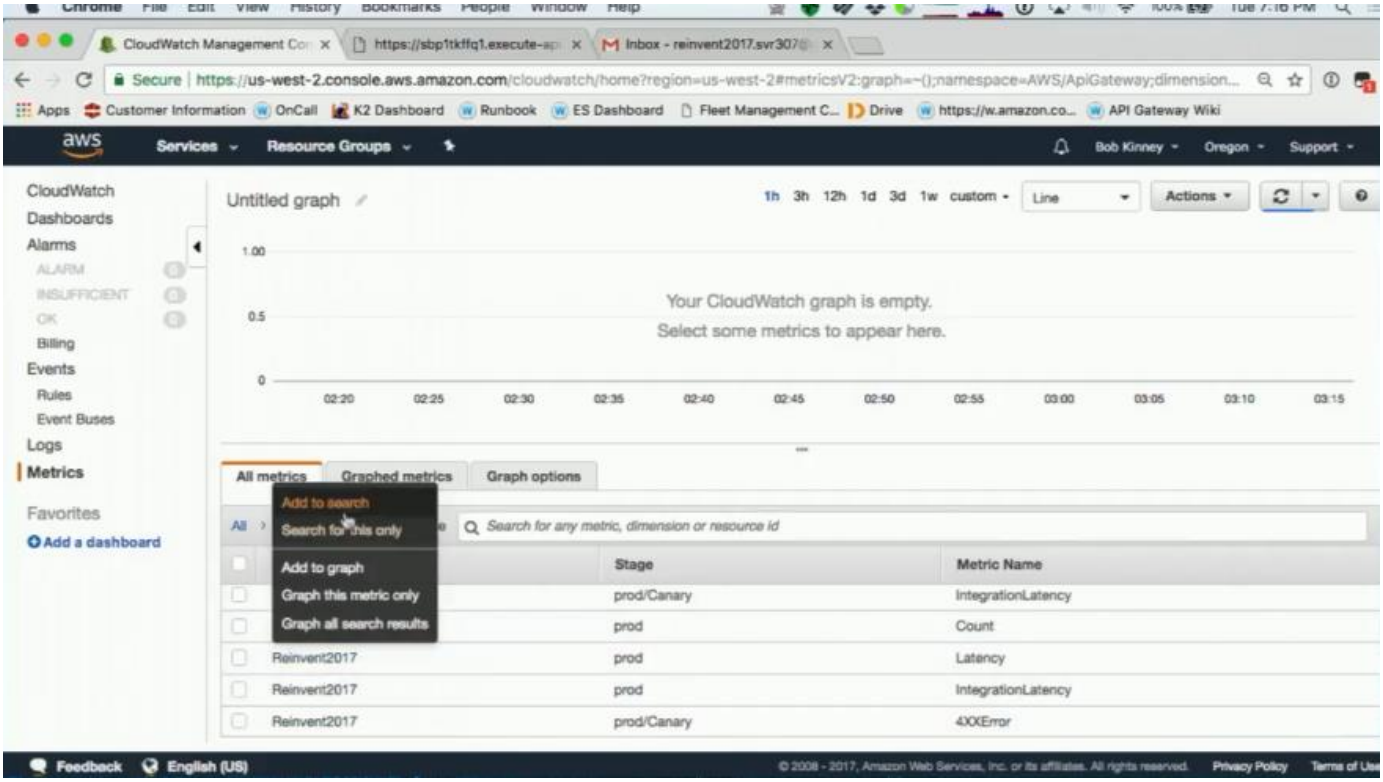

gatling-charts-highcharts-bundle-2.3.0 — grep -Ev WARN — 127×30
2017-11-28 19:15:54 5s elapsed
---- Requests
> Global (OK=0 KO=181)
> request_1 (OK=0 KO=181)
---- Errors
> status.find.in(200,304,201,202,203,204,205,206,207,208,209), b 181 (100.0%)
ut actually found 500
---- Scenario Name
[] 0%
waiting: 0 / active: 10 / done:0
2017-11-28 19:15:59 10s elapsed
---- Requests
> Global (OK=0 KO=749)
> request_1 (OK=0 KO=749)
---- Errors
> status.find.in(200,304,201,202,203,204,205,206,207,208,209), b 749 (100.0%)
ut actually found 500
---- Scenario Name
[] 0%
waiting: 0 / active: 10 / done:0

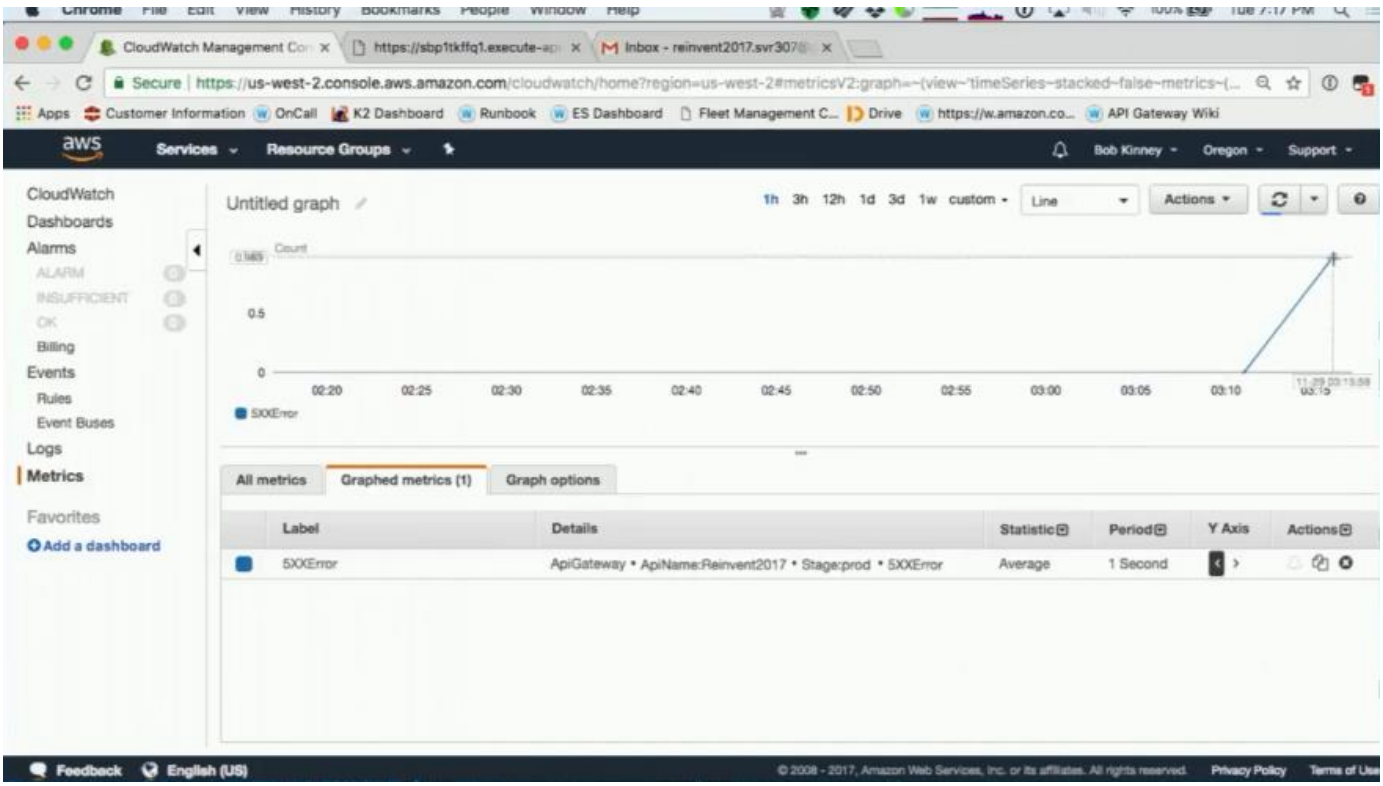

Chrome File Edit View History Bookmarks People Window Help
CloudWatch Management Con
https://sbp1tkffq1.execute-ap
Inbox - reinvent2017.svr307@
Secure | https://us-west-2.console.aws.amazon.com/cloudwatch/home?region=us-west-2
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
Metric Summary
Amazon CloudWatch monitors operational and performance metrics for your AWS cloud resources and applications. You currently have 144 CloudWatch metrics available in the US West (Oregon) region.
Browse or search your metrics to get started graphing data and creating alarms.
Browse Metrics
Search Metrics
Additional Info
Getting Started Guide
Monitoring Scripts Guide
Overview and Features
Documentation
Forums
Report an Issue
Alarm Summary
You do not have any alarms created in the US West (Oregon) region. Alarms allow you to send notifications or execute AutoScaling actions in response to any CloudWatch metric.
You can now use Amazon CloudWatch alarms to monitor the estimated charges on your AWS bill and receive email alerts whenever charges exceed a threshold you define. Visit the CloudWatch US East (N. Virginia) region to manage your billing alarms.
Go to CloudWatch US East (N. Virginia) region
Service Health
Current Status
Details
Amazon CloudWatch Service
Service is operating normally
View complete service health details
Feedback English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use


Chrome File Edit View History Bookmarks People Window Help
CloudWatch Management Con
https://sbp1tkffq1.execute-ap
Inbox - reinvent2017.svr307@
Secure | https://us-west-2.console.aws.amazon.com/cloudwatch/home?region=us-west-2#metricsV2:
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
Untitled graph
1h 3h 12h 1d 3d 1w custom
Line
Actions
1.00
0.5
0
Your CloudWatch graph is empty.
Select some metrics to appear here.
02:20 02:25 02:30 02:35 02:40 02:45 02:50 02:55 03:00 03:05 03:10 03:15
All metrics
Graphed metrics
Graph options
Search for any metric, dimension or resource id
144 Metrics
ApiGateway
38 Metrics
DynamoDB
4 Metrics
Lambda
36 Metrics
Logs
18 Metrics
S3
8 Metrics
SNS
8 Metrics
Feedback English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

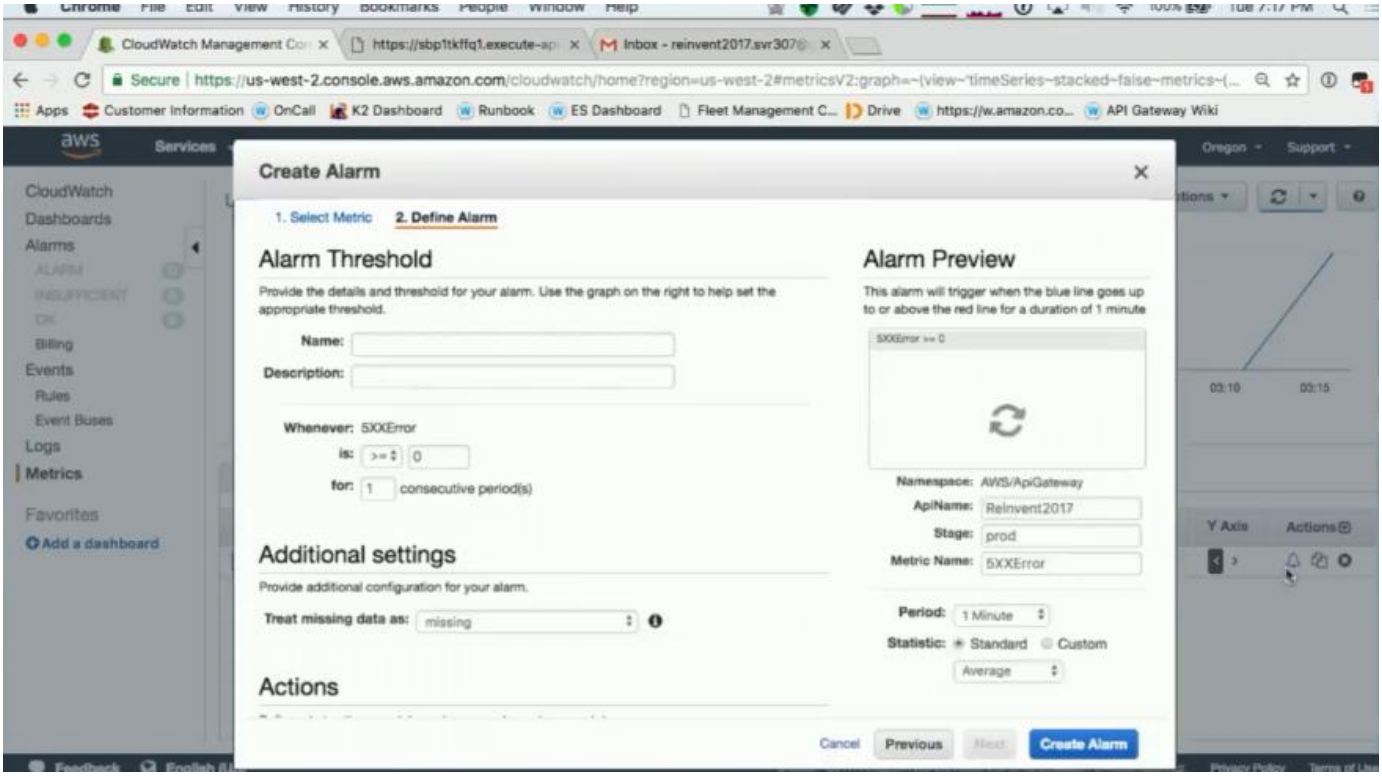

Chrome File Edit View History Bookmarks People Window Help
CloudWatch Management Con
https://sbp1tkffq1.execute-ap
Inbox - reinvent2017.svr307
Secure | https://us-west-2.console.aws.amazon.com/cloudwatch/home?region=us-west-2#metricsV2:graph=~();namespace=AWS/ApiGateway
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
Untitled graph
1h 3h 12h 1d 3d 1w custom
Line
Actions
1.00
0.5
0
Your CloudWatch graph is empty.
Select some metrics to appear here.
02:20 02:25 02:30 02:35 02:40 02:45 02:50 02:55 03:00 03:05 03:10 03:15
All metrics
Graphed metrics
Graph options
All > ApiGateway
Search for any metric, dimension or resource id
38 Metrics
By Method
5 Metrics
By Stage
19 Metrics
By Api Name
14 Metrics
Feedback English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

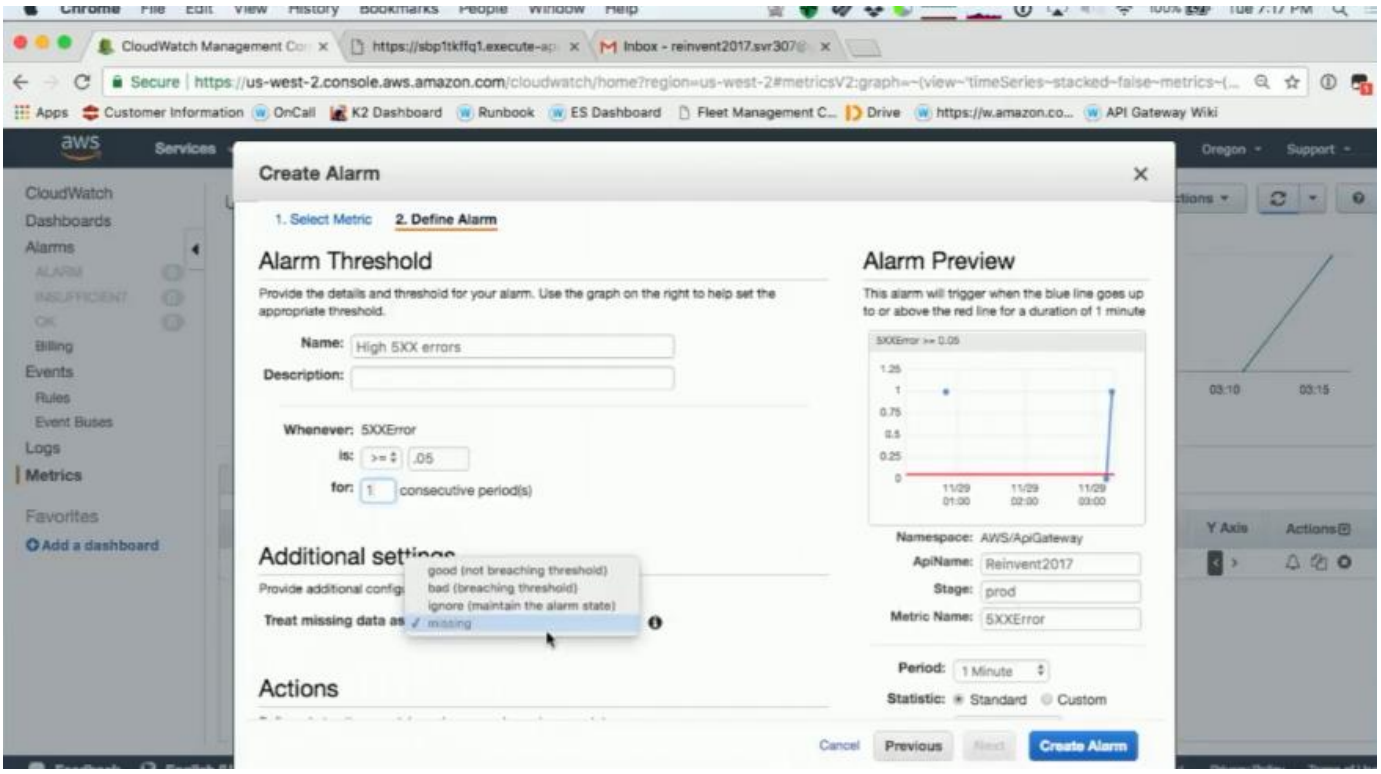

Chrome File Edit View History Bookmarks People Window Help
CloudWatch Management Con
https://sbp1tkffq1.execute-ap
Inbox - reinvent2017.svr307
Secure | https://us-west-2.console.aws.amazon.com/cloudwatch/home?region=us-west-2#metricsV2:graph=~();namespace=AWS/ApiGateway;dimension...
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
Untitled graph
1h 3h 12h 1d 3d 1w custom
Line
Actions
1.00
0.5
0
Your CloudWatch graph is empty.
Select some metrics to appear here.
02:20 02:25 02:30 02:35 02:40 02:45 02:50 02:55 03:00 03:05 03:10 03:15
All metrics
Graphed metrics
Graph options
All > ApiGateway > By Stage
Search for any metric, dimension or resource id
ApiName (19)
Stage
Metric Name
Reinvent2016 prod Latency
Reinvent2016 prod Count
Reinvent2016 prod IntegrationLatency
Reinvent2016 prod 4XXError
Reinvent2016 prod 5XXError
Feedback English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use



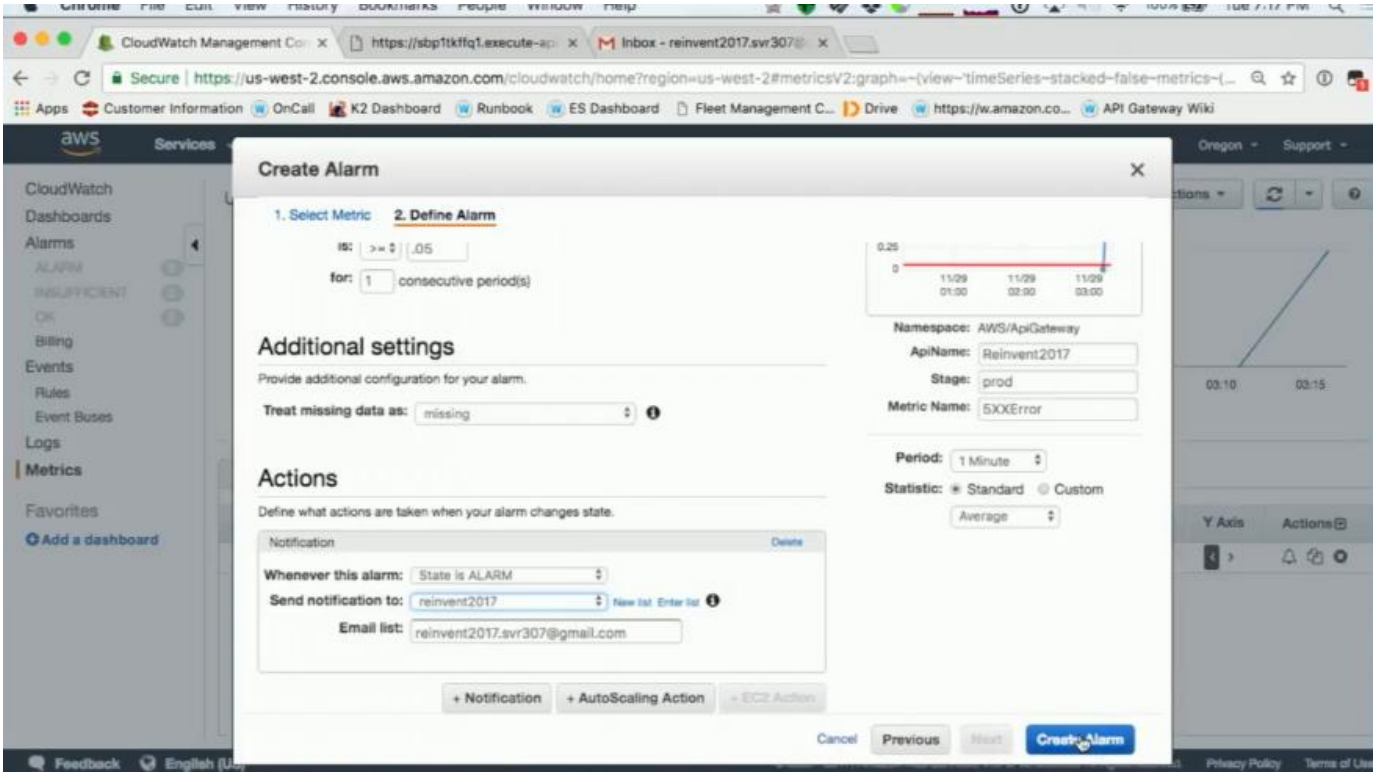


We can see the 500 error count as above. We can then set an alarm by clicking on the bell icon as below

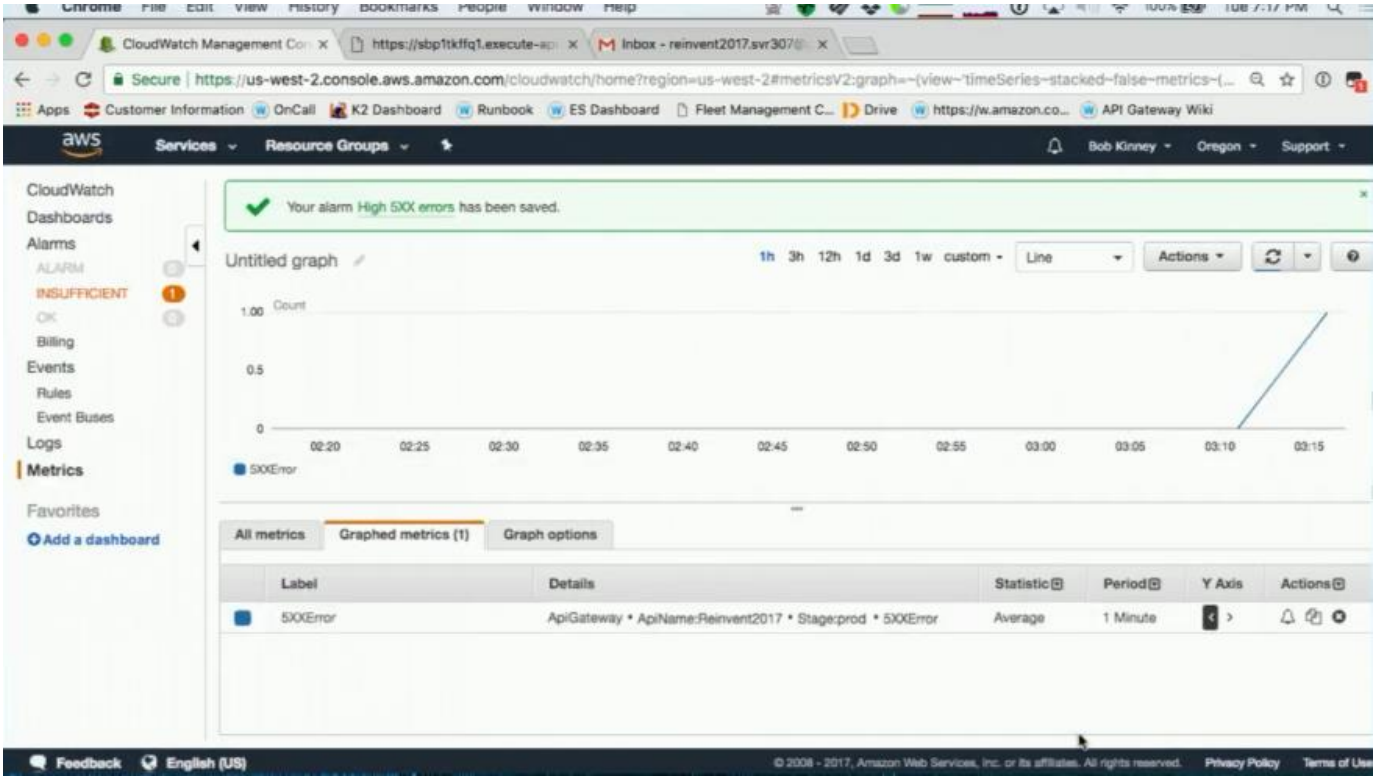


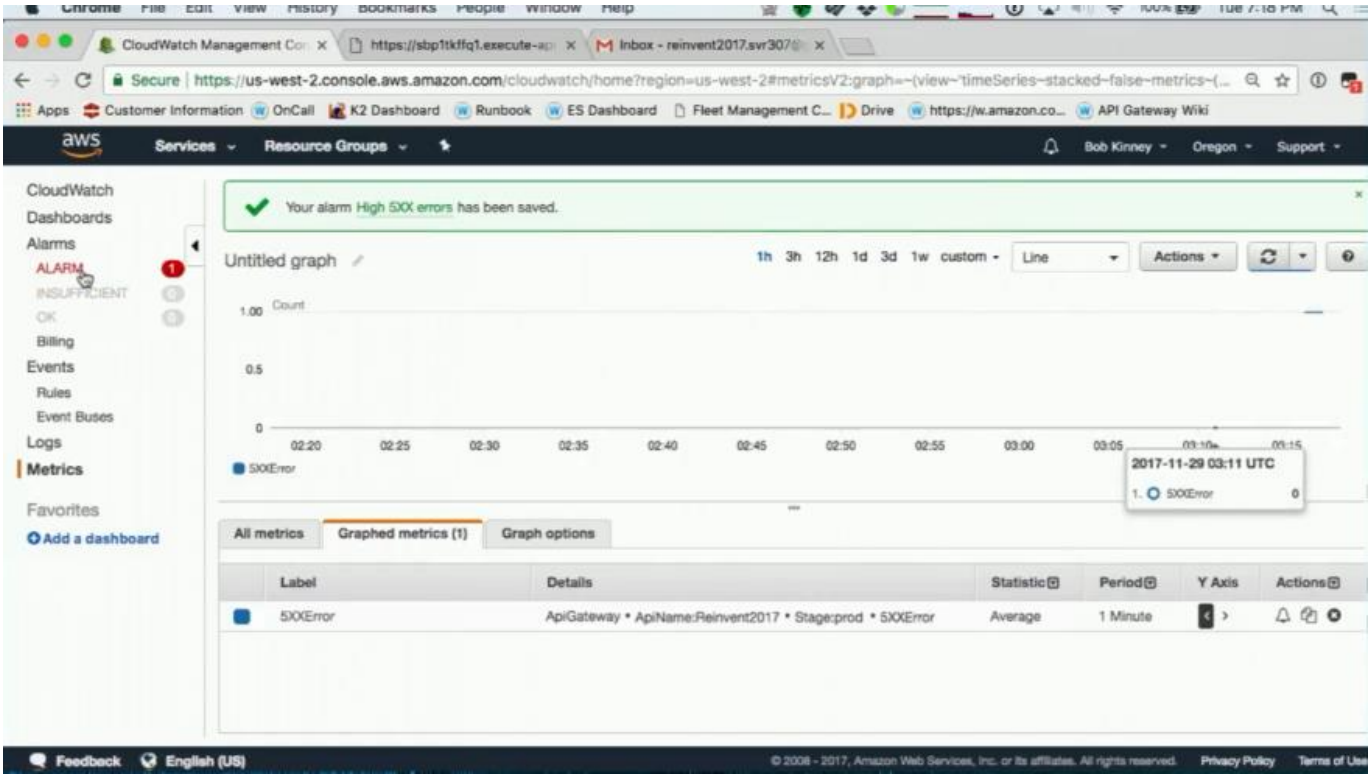


We are going to create an alarm for getting 500 responses 50% of the calls within a duration

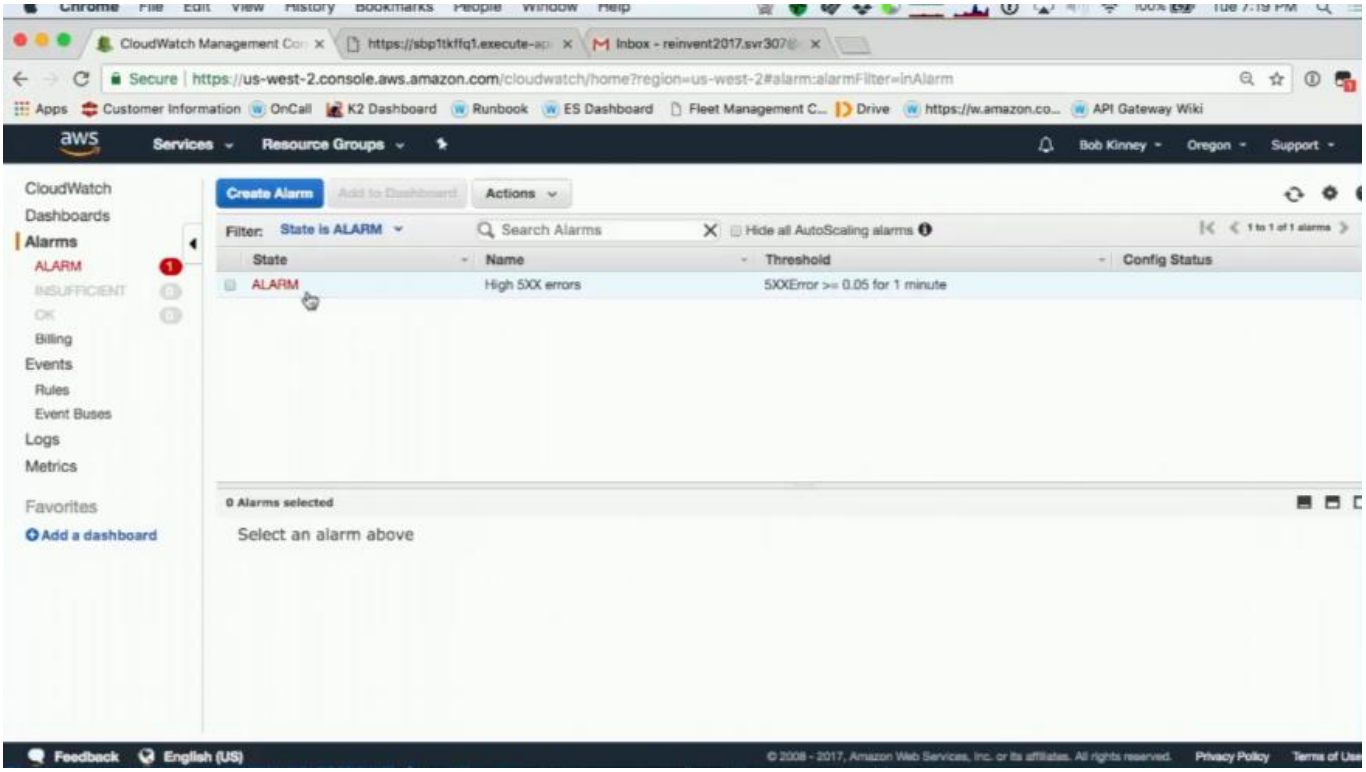


We then supply an email address to have the alarm notification sent to

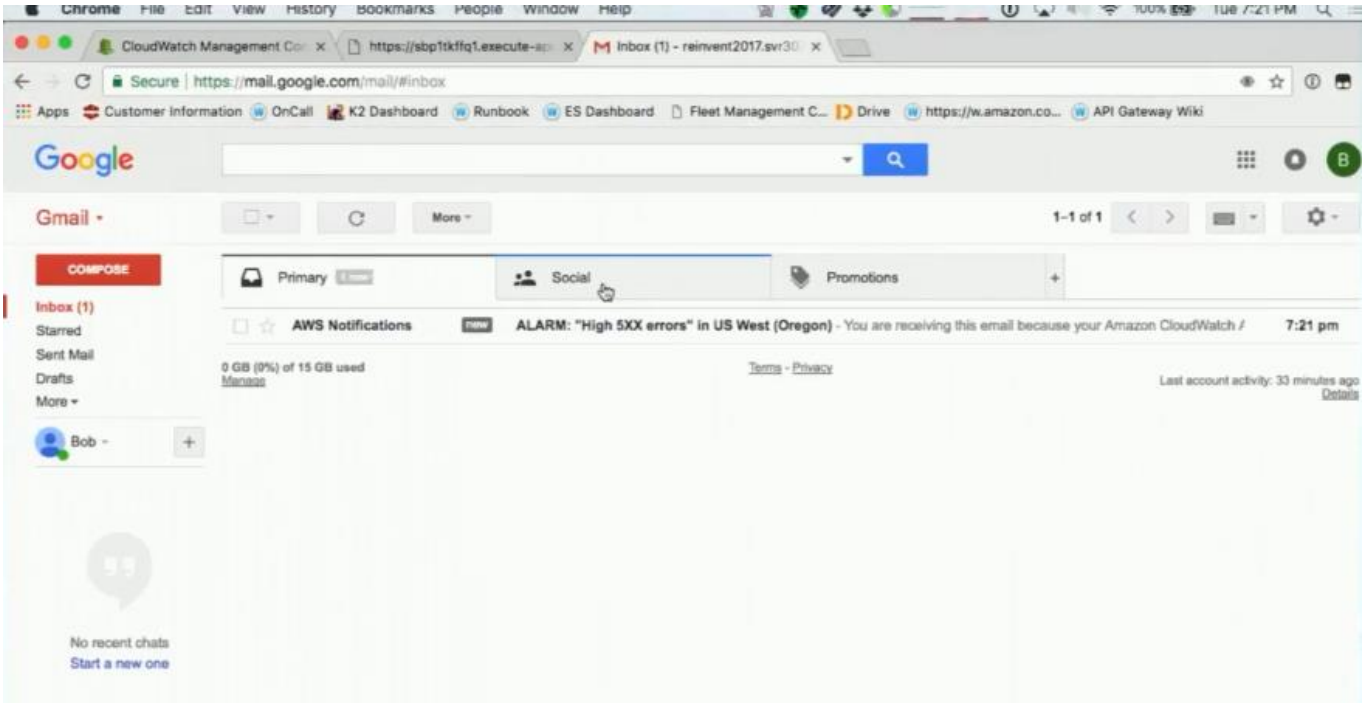




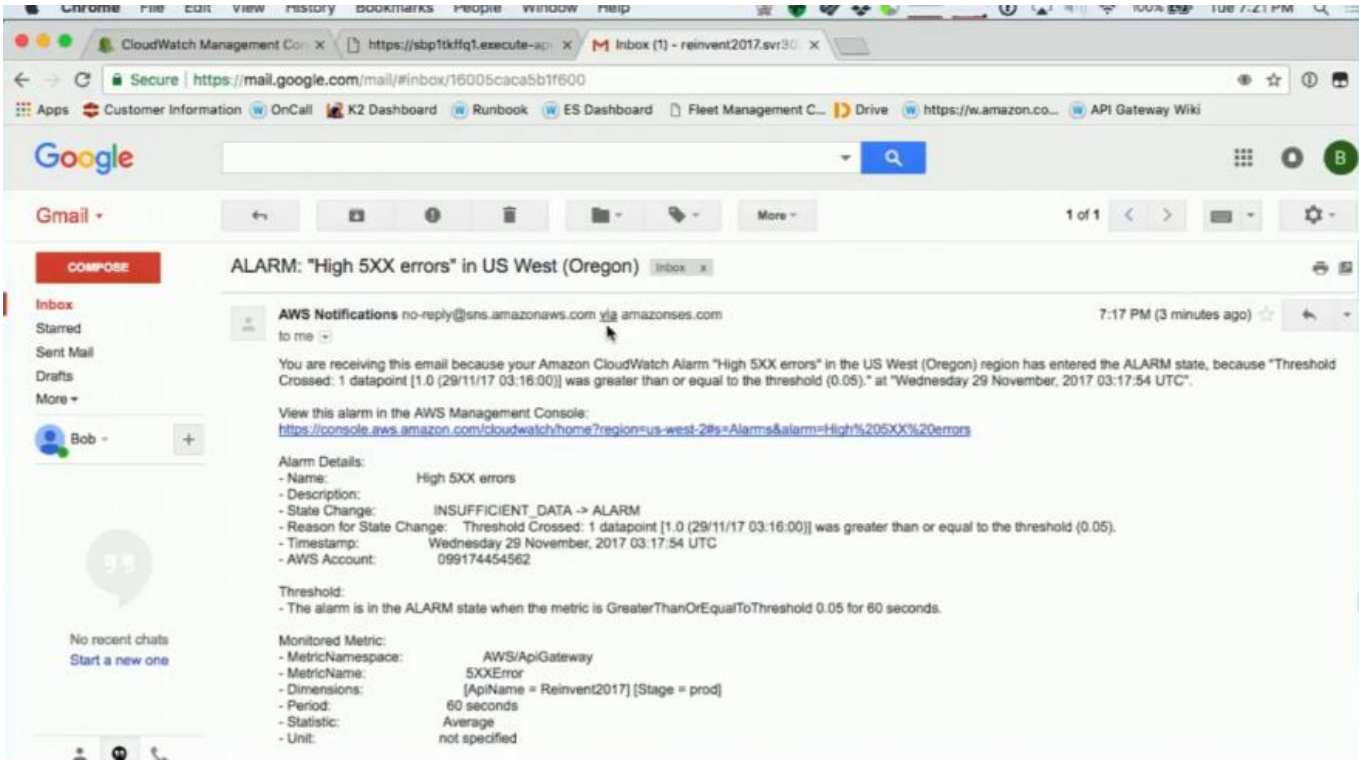
We can then see an alarm and can check our email address.



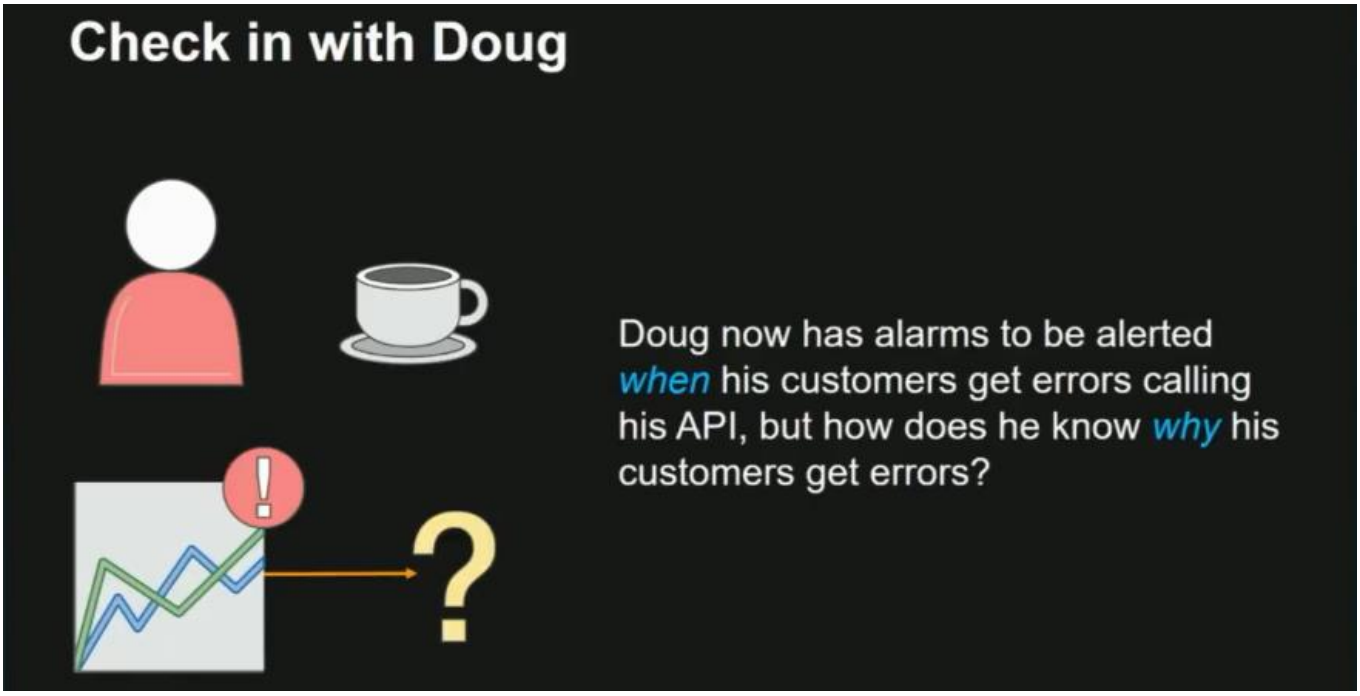
We are now in an alarmed state



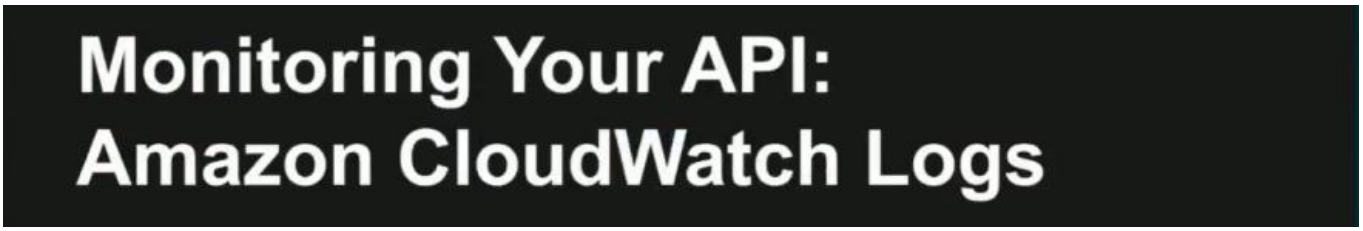
And we can see the email sent to us about the alarm from CloudWatch



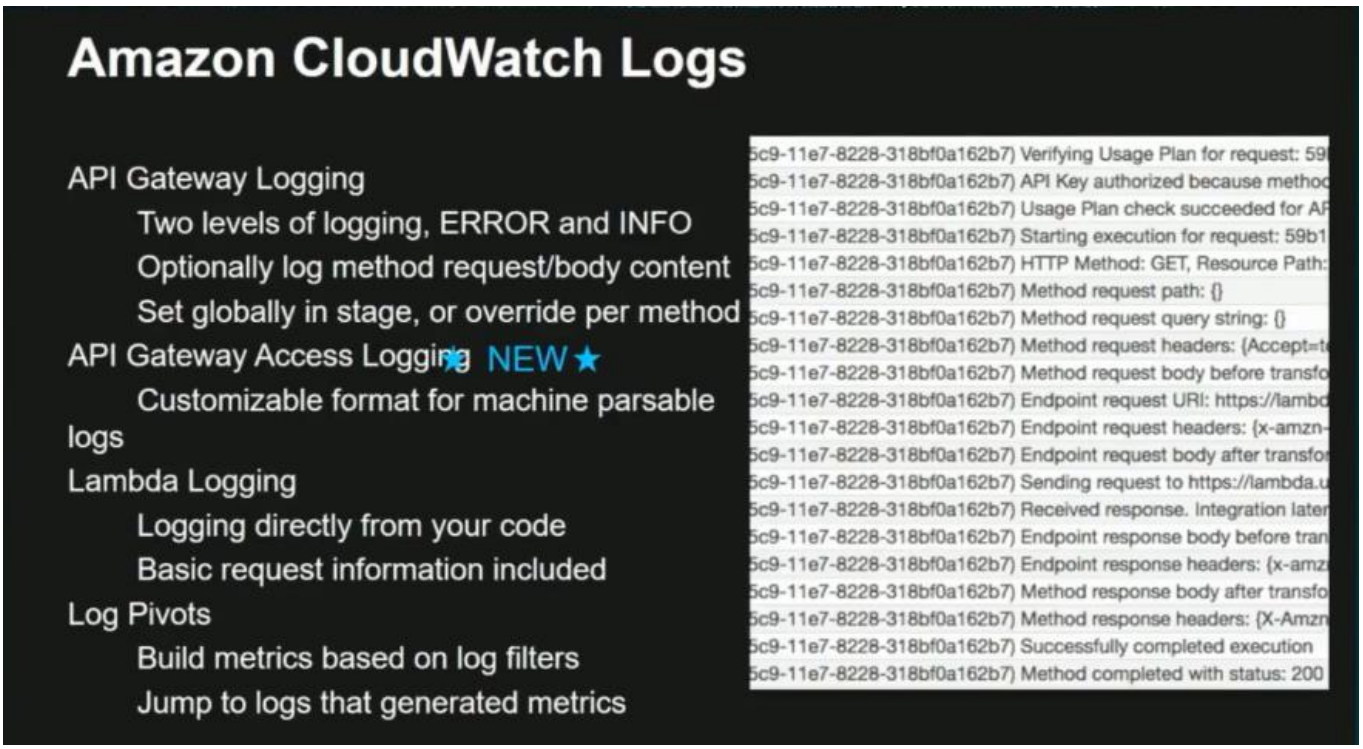
Demo



We now know when we are having problems in our APIs but we don't yet know why we are having those errors.



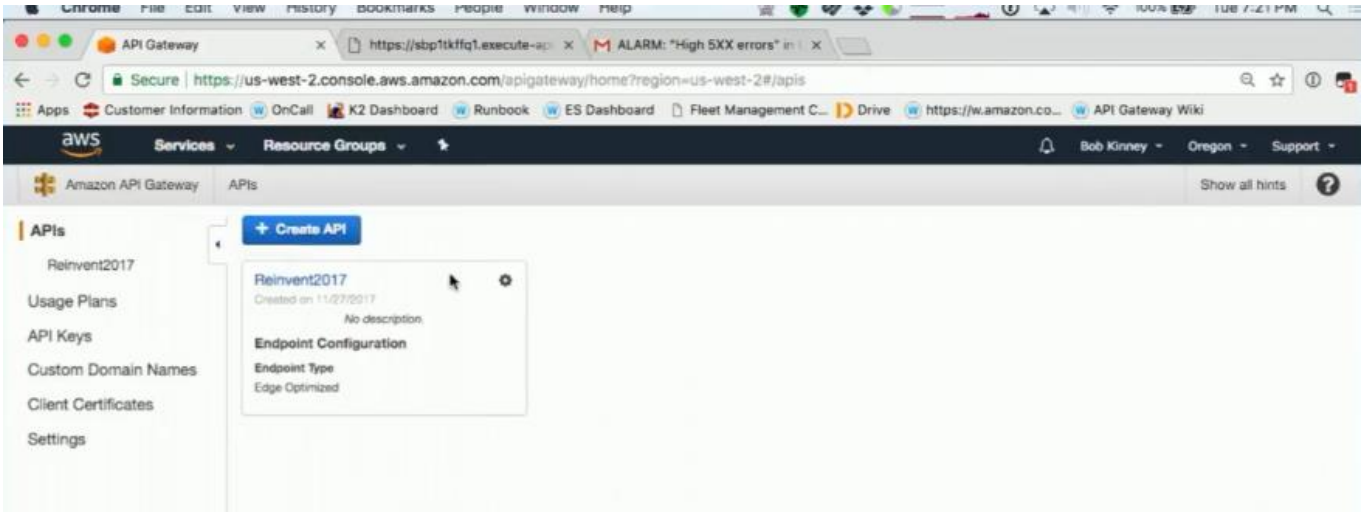
This talks about logging using Amazon CloudWatch Logs.

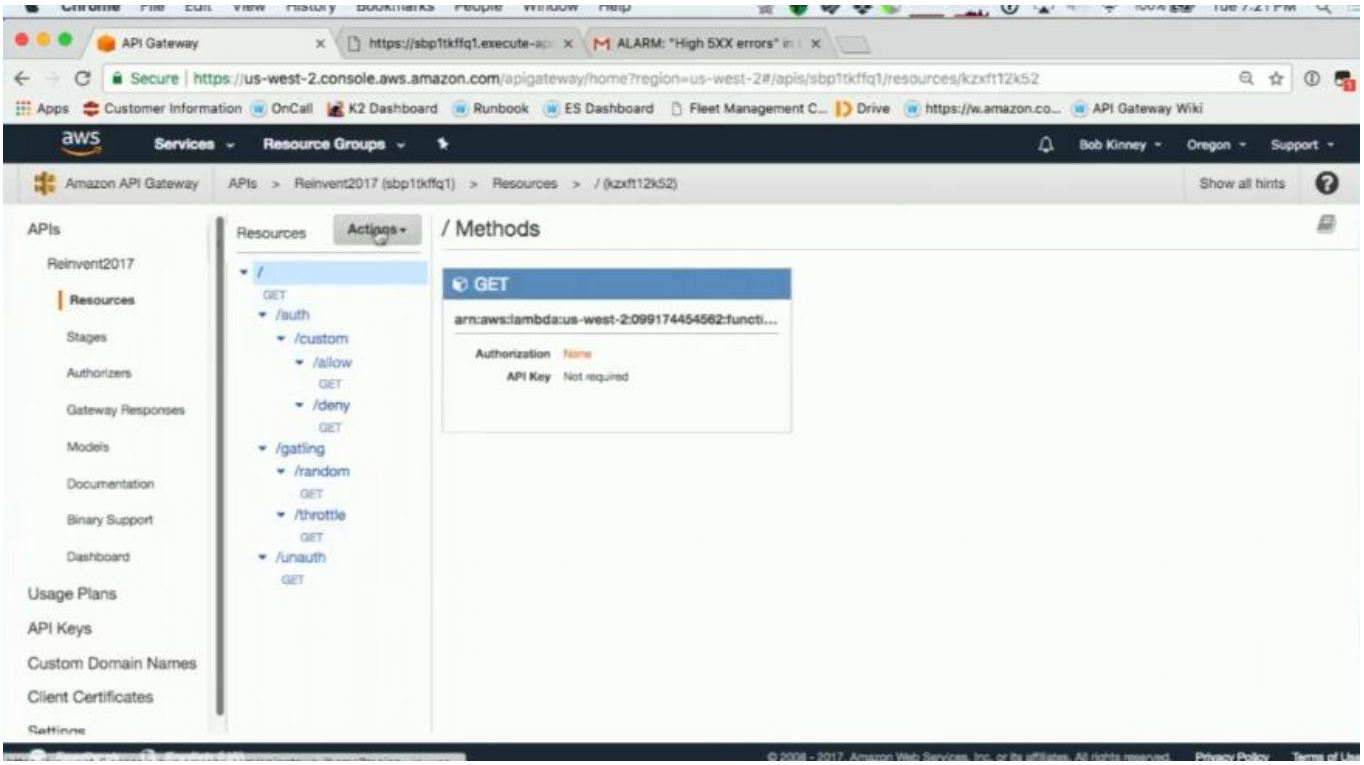


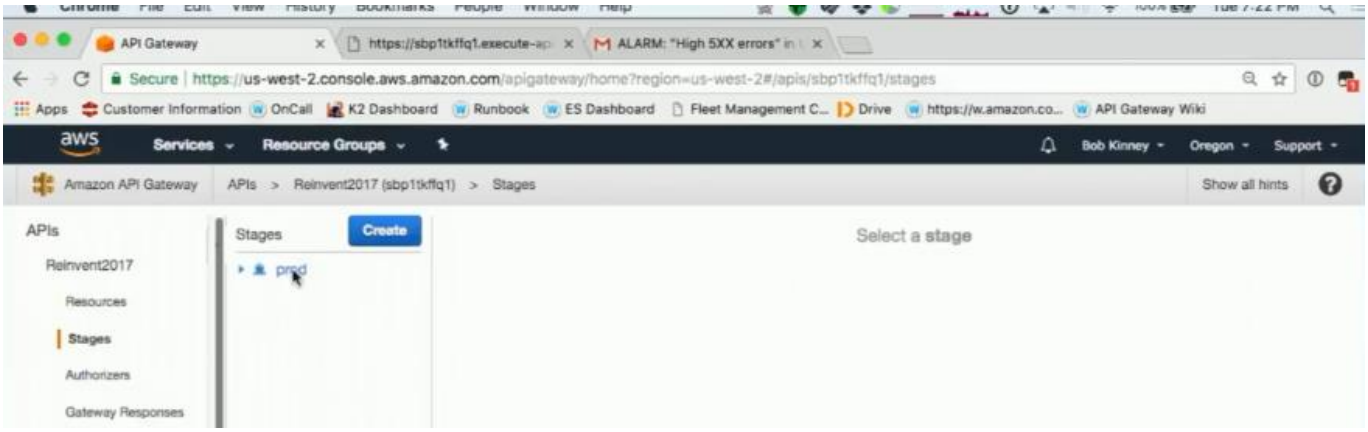
API Gateway offers 2 basic logging types. ***Logging logs*** for the basic request/response, API transformation logging and ***Access logs***, ***Lambda Logs*** for console.logs within your lambda functions.

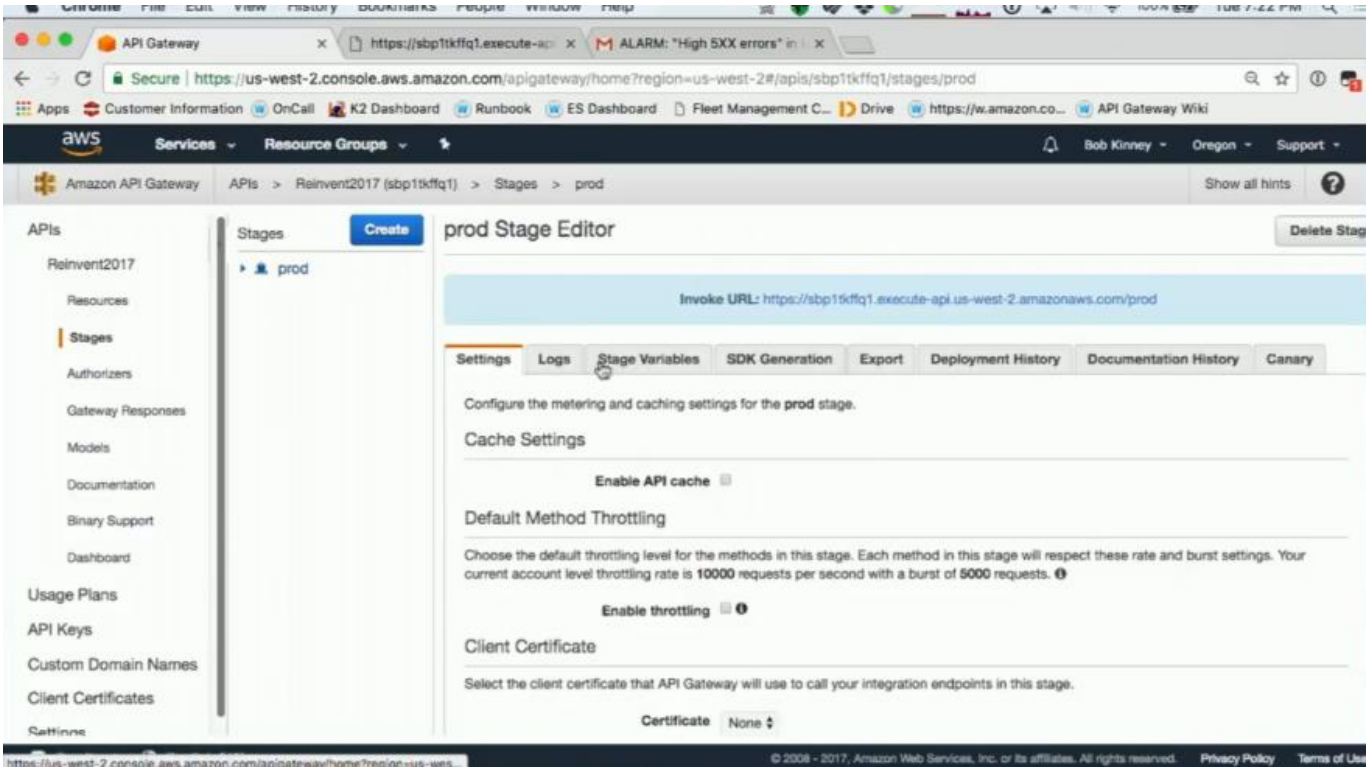
Demo

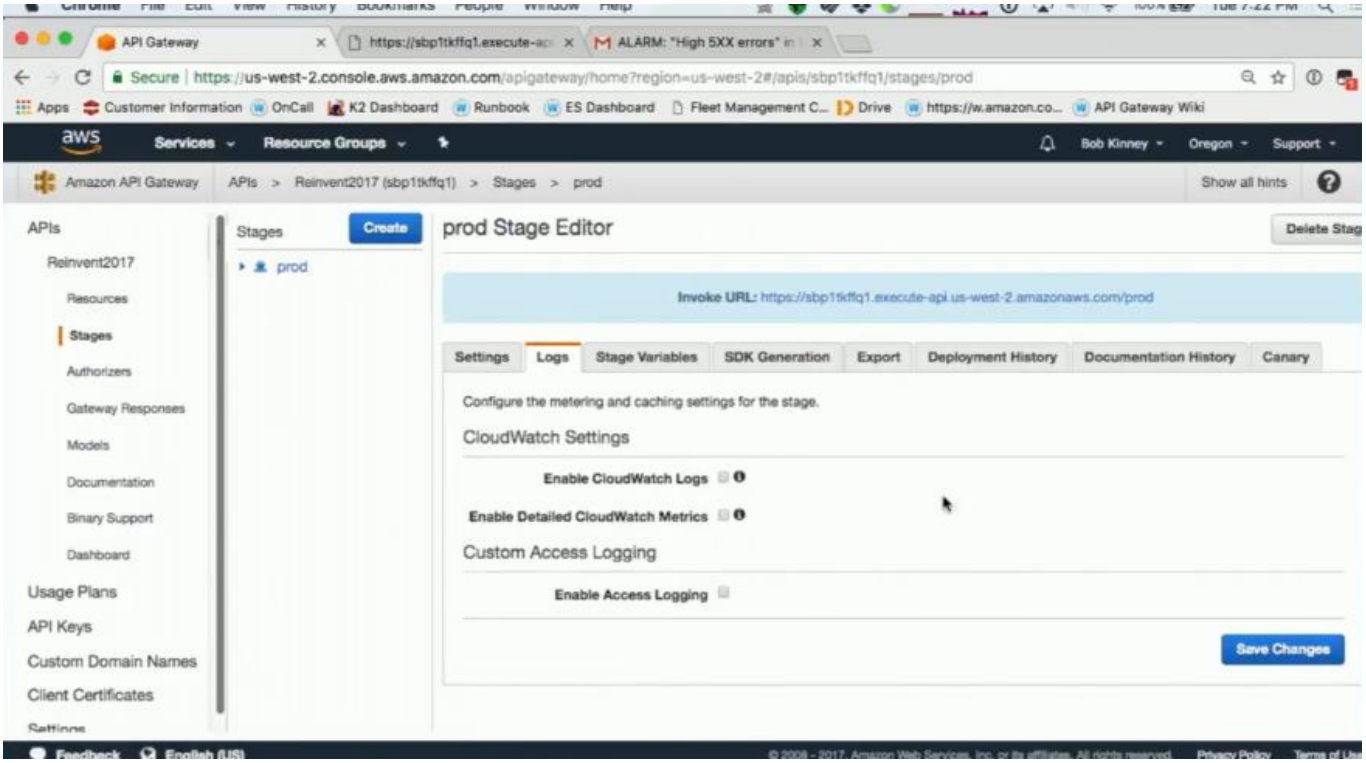
Let us now try to diagnose what exactly is going on with our API.

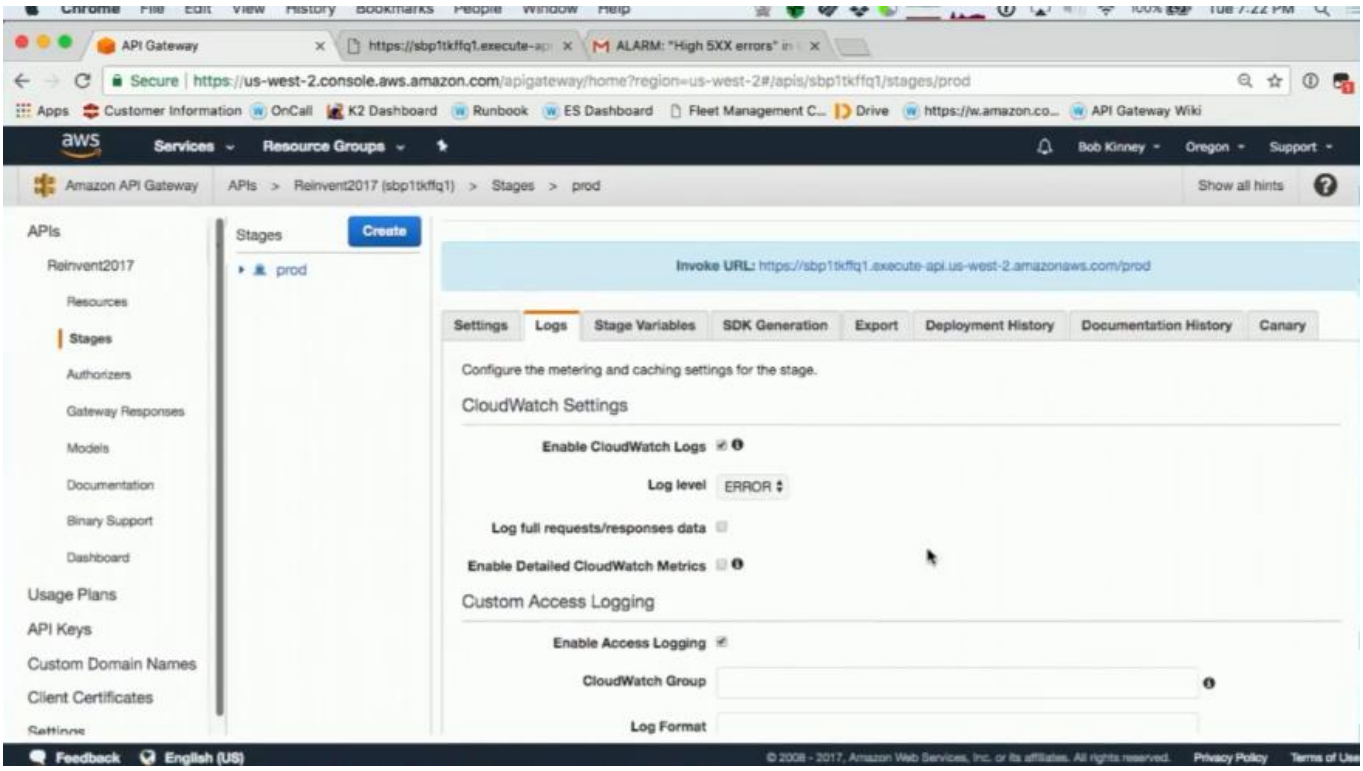




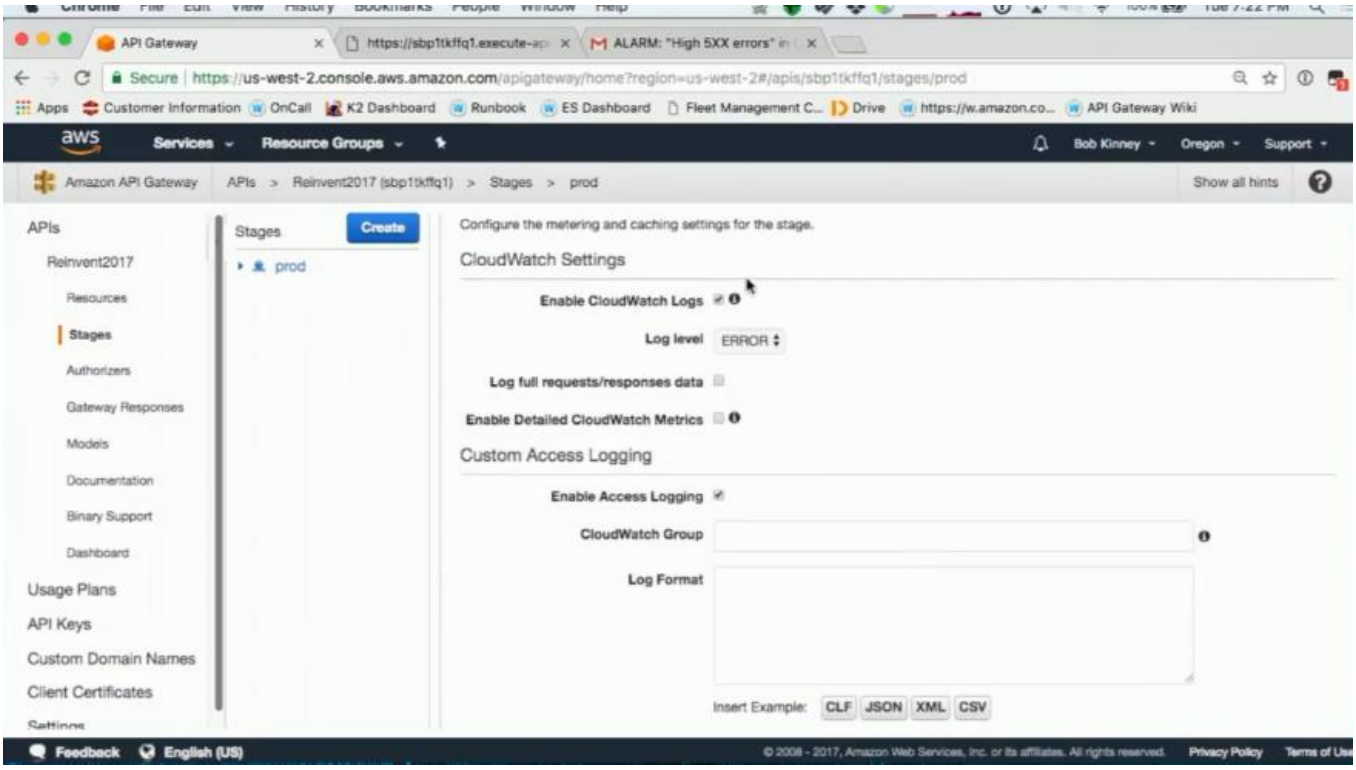

API Gateway
ALARM: "High 5XX errors"
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/stages
Apps
Customer Information
OnCall
K2 Dashboard
Runbook
ES Dashboard
Fleet Management C...
Drive
https://w.amazon.co...
API Gateway Wiki
aws
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Stages
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Stages
Create
prod
Select a stage


API Gateway
ALARM: "High 5XX errors"
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/stages/prod
Apps
Customer Information
OnCall
K2 Dashboard
Runbook
ES Dashboard
Fleet Management C...
Drive
https://w.amazon.co...
API Gateway Wiki
aws
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Stages > prod
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Stages
Create
prod
prod Stage Editor
Delete Stag
Invoke URL: https://sbp1tkffq1.execute-api.us-west-2.amazonaws.com/prod
Settings
Logs
Stage Variables
SDK Generation
Export
Deployment History
Documentation History
Canary
Configure the metering and caching settings for the prod stage.
Cache Settings
Enable API cache
Default Method Throttling
Choose the default throttling level for the methods in this stage. Each method in this stage will respect these rate and burst settings. Your current account level throttling rate is 10000 requests per second with a burst of 5000 requests.
Enable throttling
Client Certificate
Select the client certificate that API Gateway will use to call your integration endpoints in this stage.
Certificate
None
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

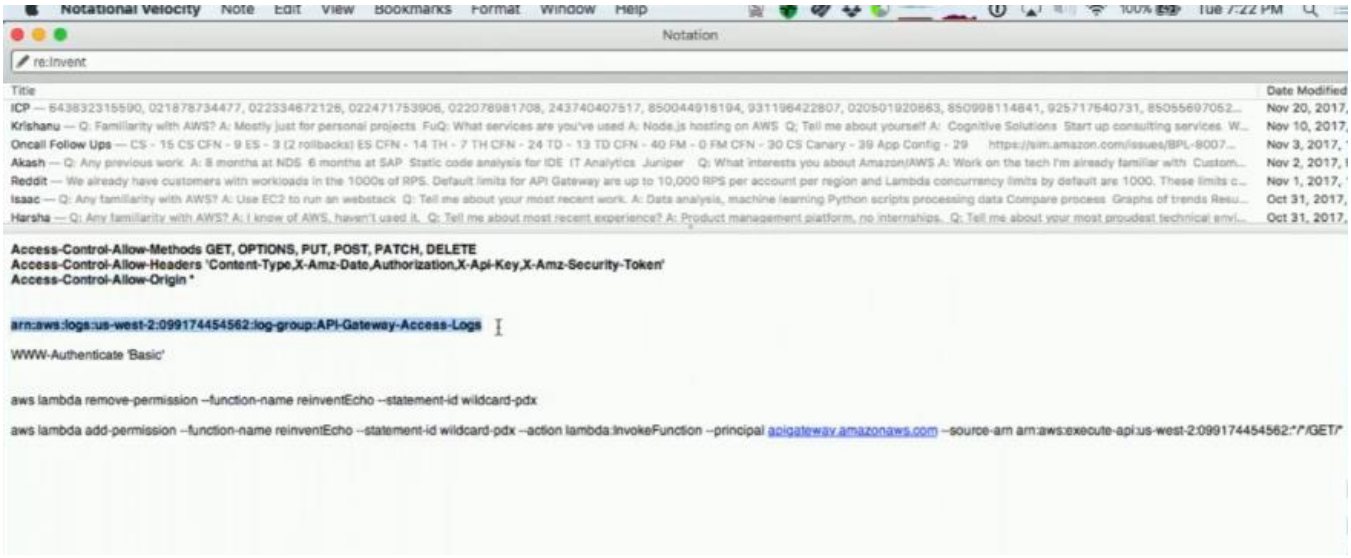


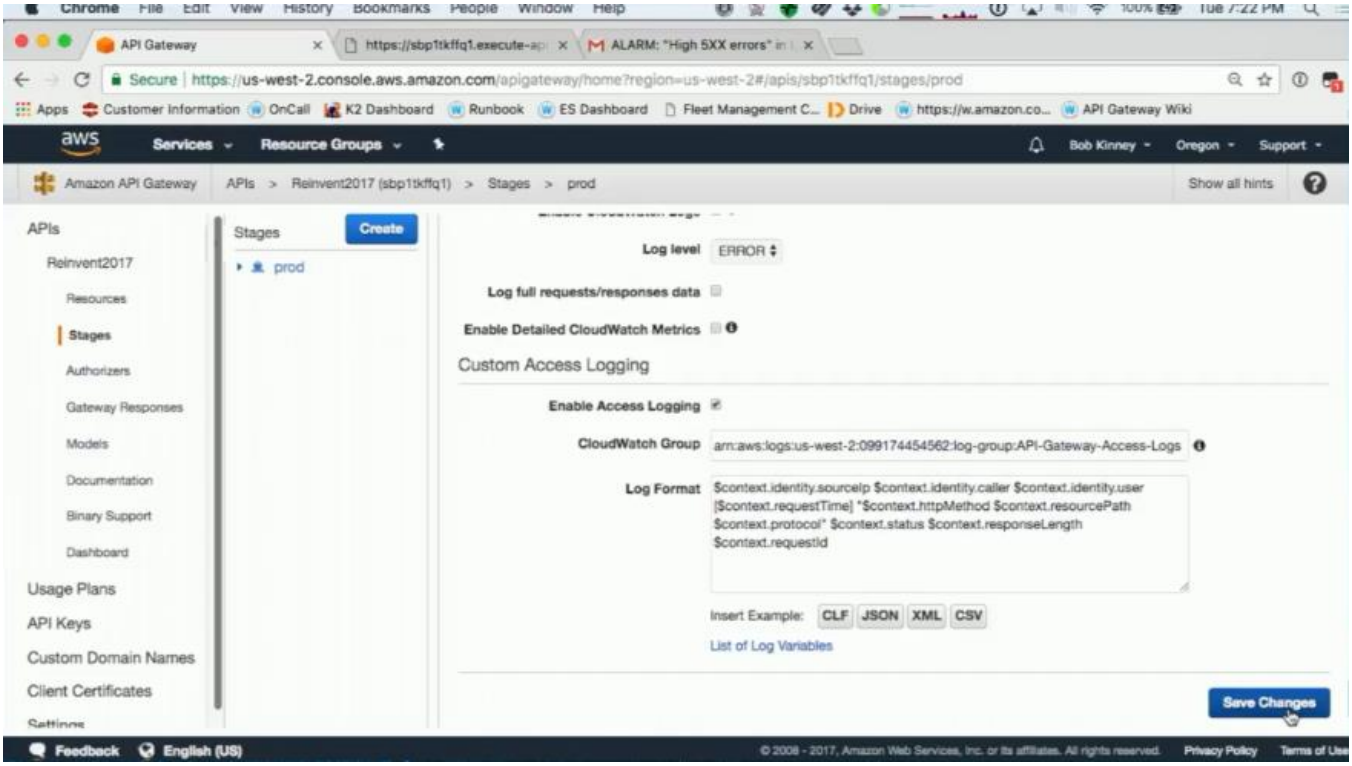


We then enabled logs at the ERROR and Access levels as above.

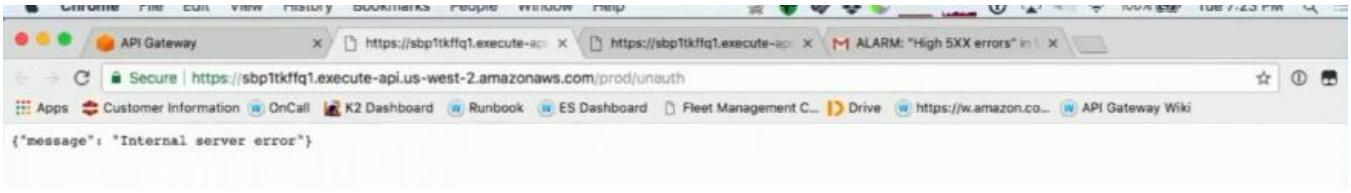


For Access logging, we can actually provide a Log Group to provide all of our Access logs to. This will allow you to aggregate all your access logs across multiple APIs if you want.

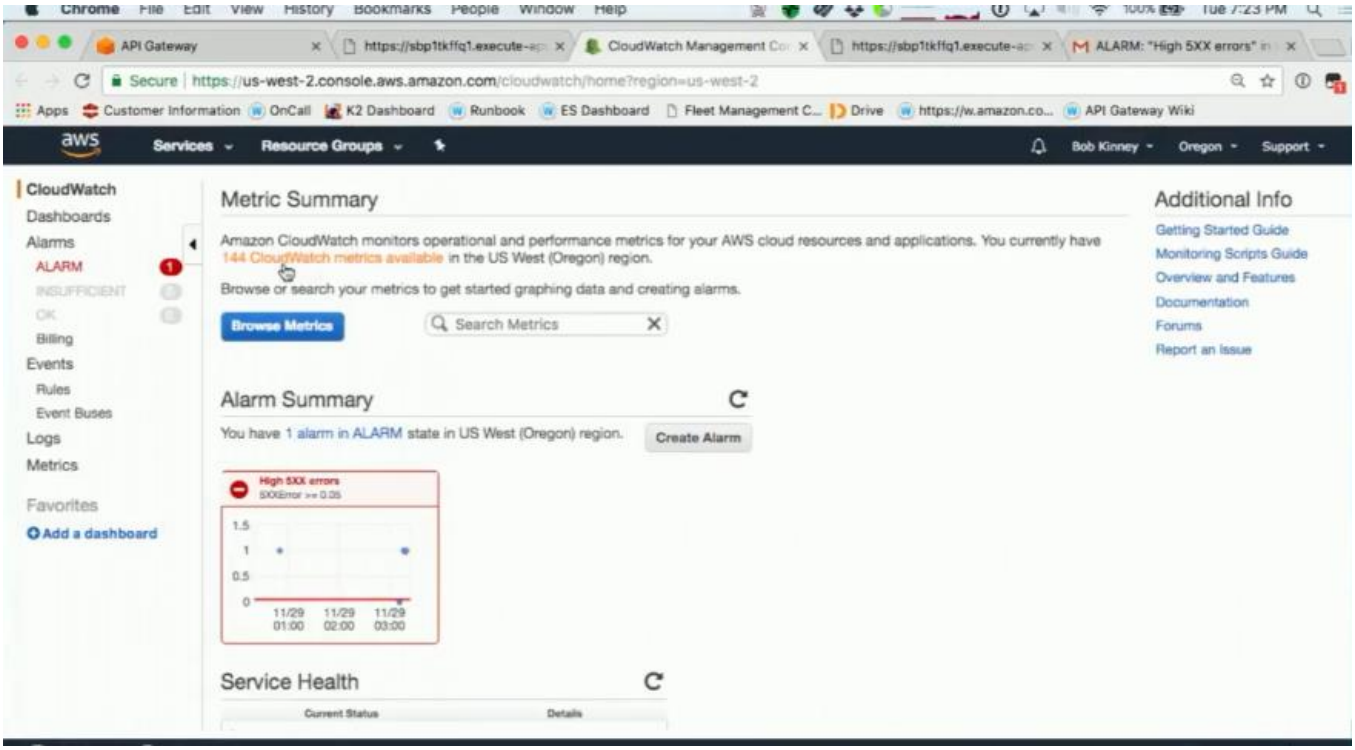




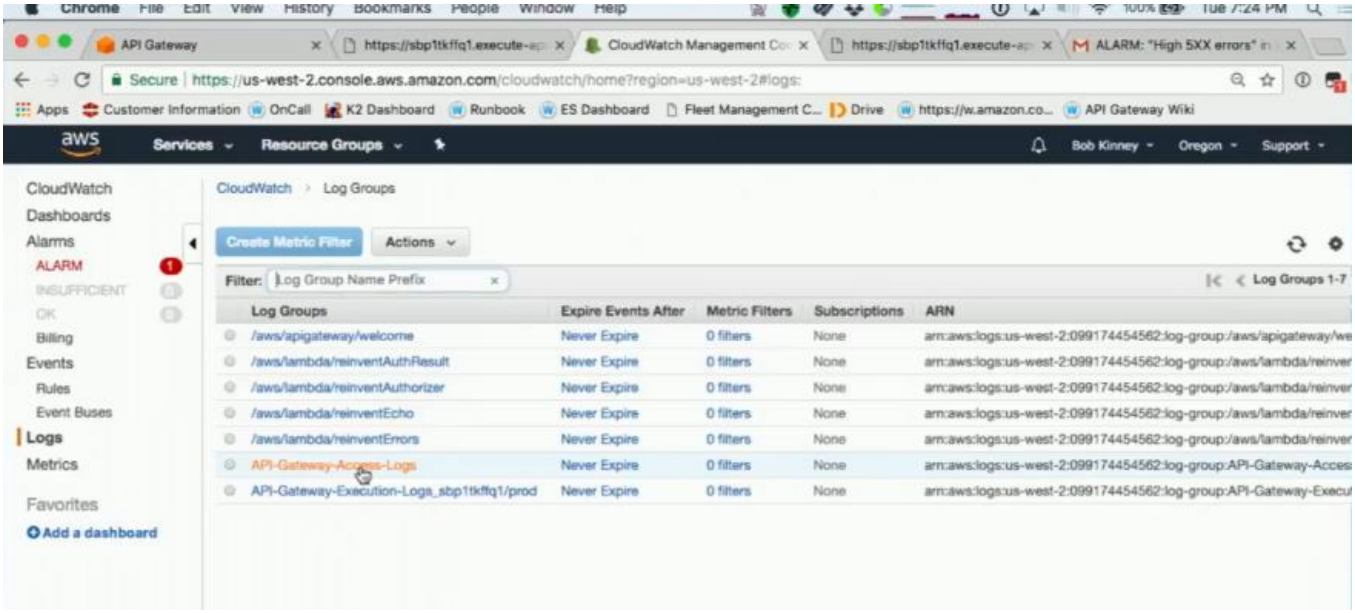
We are going to be using the CLF for the Common Log Format option as above. We then Save the changes.

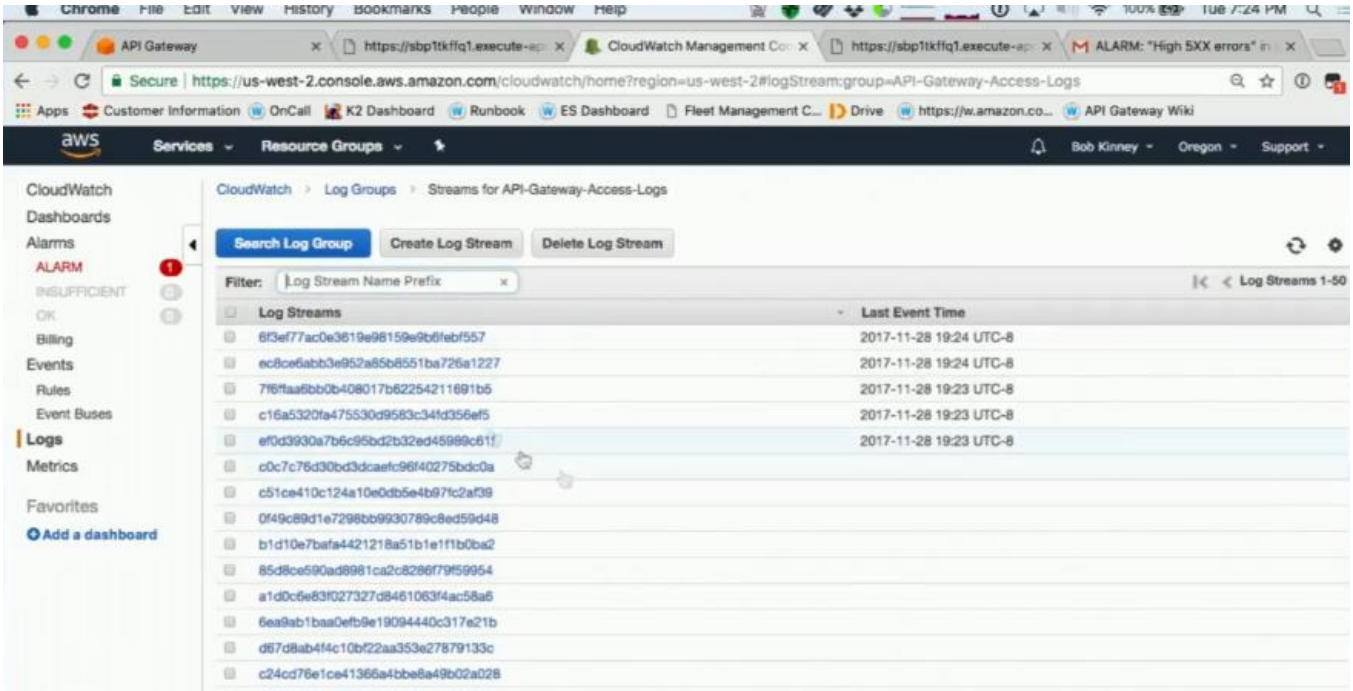


To test, we can then hit this 500 error response endpoint a couple of times

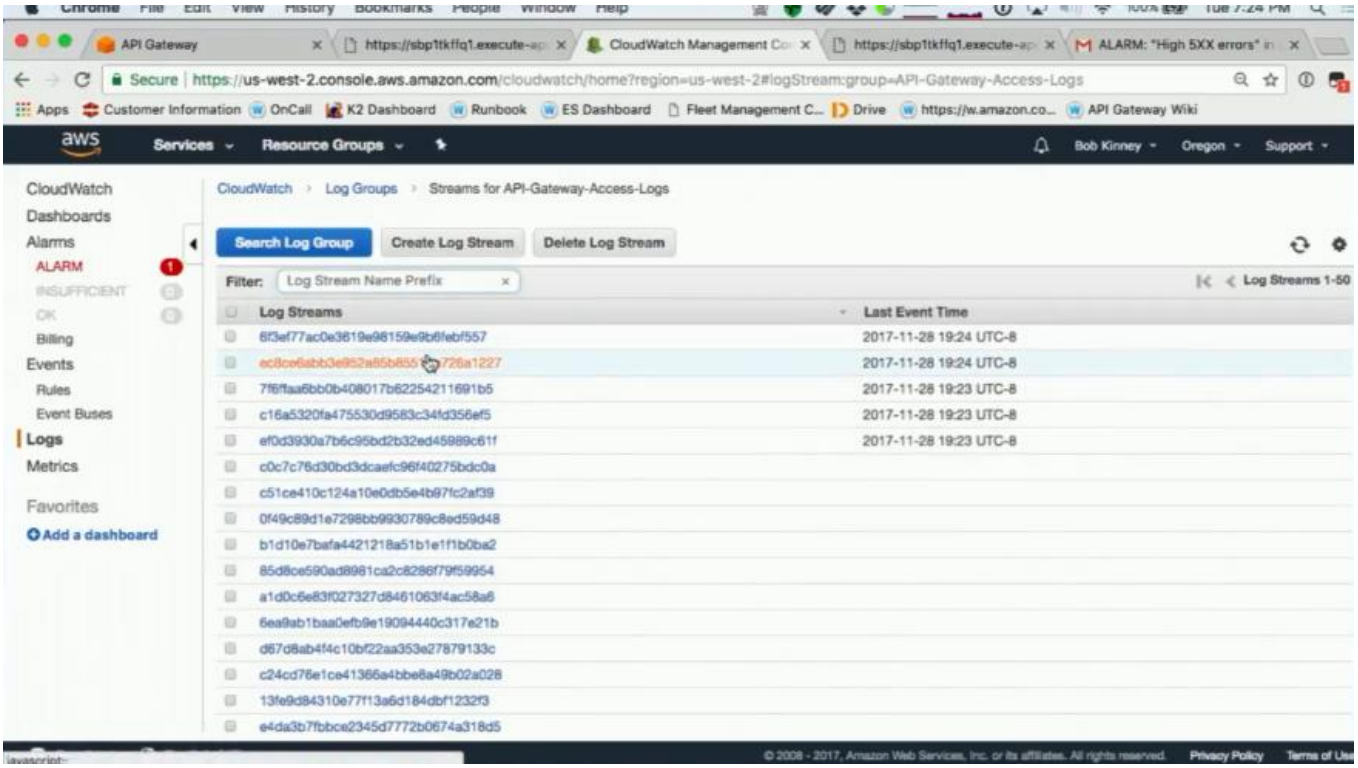


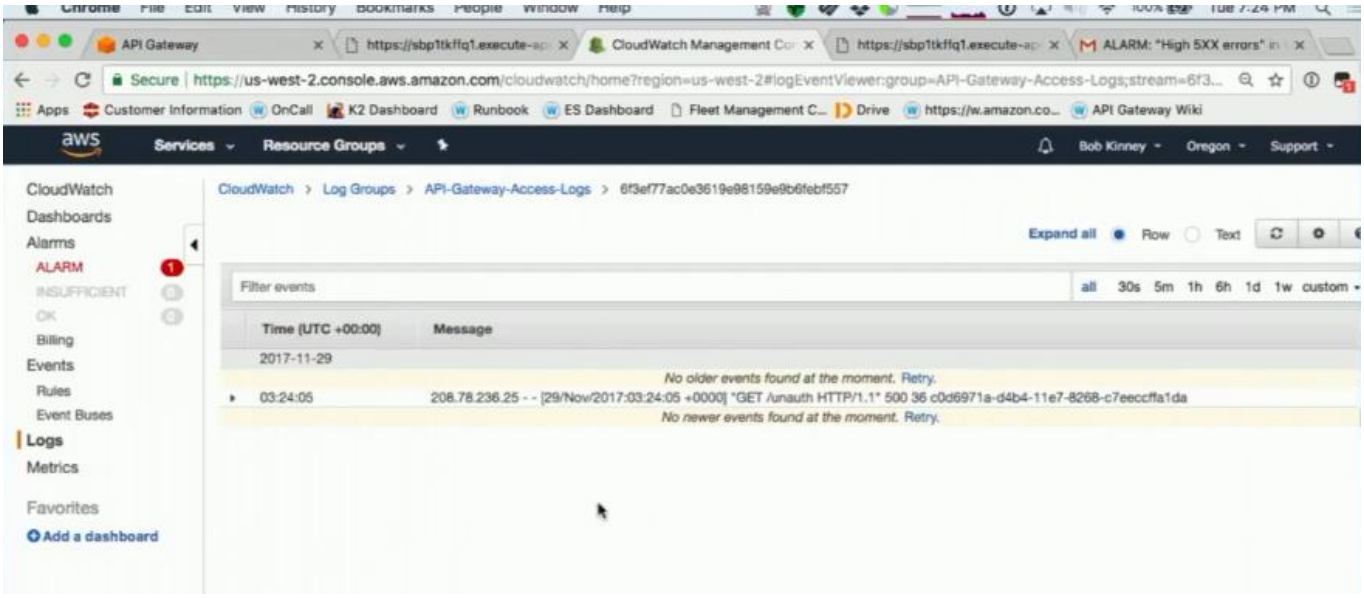
In CloudWatch, we should now be able to find and see the specific request Ids for those 500 calls inside our access logs.

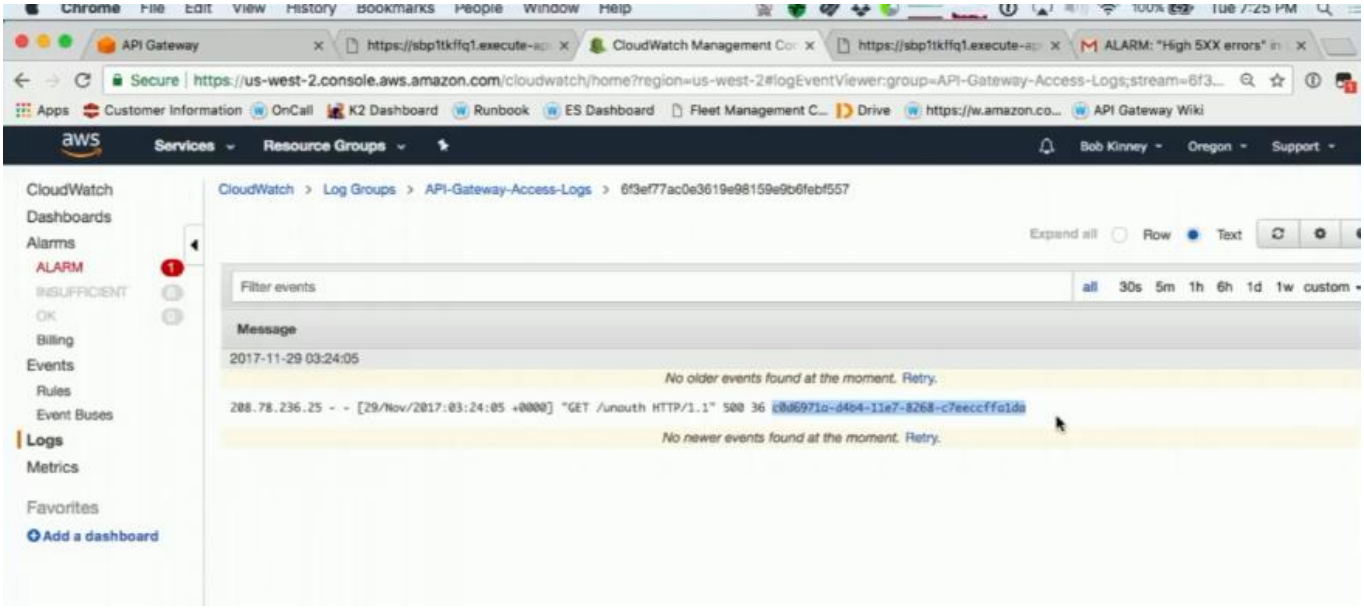




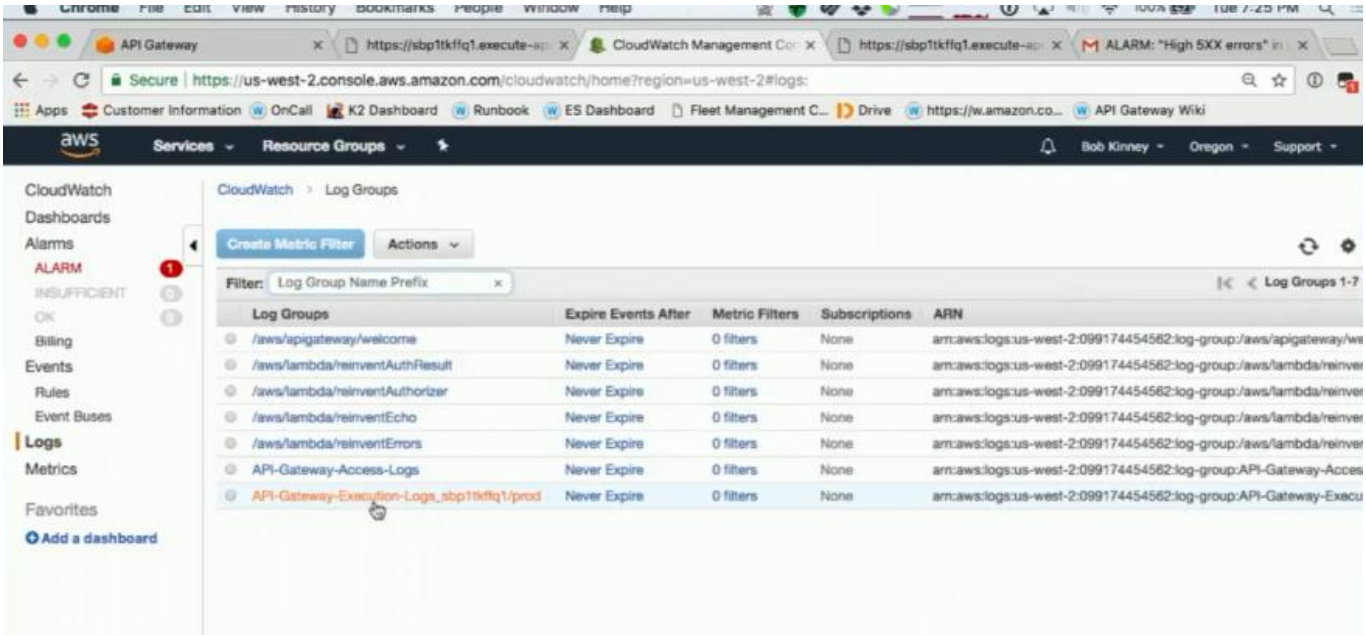
This is our Access Logs group filtered by time


Secure | https://us-west-2.console.aws.amazon.com/cloudwatch/home?region=us-west-2#logStream:group=API-Gateway-Access-Logs
Services
Resource Groups
Bob Kinney
Oregon
Support
CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
CloudWatch > Log Groups > Streams for API-Gateway-Access-Logs
Search Log Group
Create Log Stream
Delete Log Stream
Filter: Log Stream Name Prefix
Log Streams 1-50
Log Streams
Last Event Time
6f3ef77ac0e3619e98159e9b6febf557 2017-11-28 19:24 UTC-8
ec8ce6abb3e952a85b8551 726a1227 2017-11-28 19:24 UTC-8
7f6ffaa6bb0b408017b6225421169 1b5 2017-11-28 19:23 UTC-8
c16a5320fa475530d9583c34fd356ef5 2017-11-28 19:23 UTC-8
ef0d3930a7b6c95bd2b32ed45989c61f 2017-11-28 19:23 UTC-8
c0c7c76d30bd3dcaefc96f40275bdc0a
c51ce410c124a10e0db5e4b97fc2af39
0f49c89d1e7298bb9930789c8ed59d48
b1d10e7bafa4421218a51b1e1f1b0ba2
85d8ce590ad8981ca2c8286f79f59954
a1d0c6e83f027327d8461063f4ac58a6
6ea9ab1baa0efb9e19094440c317e21b
d67d8ab4f4c10bf22aa353e27879133c
c24cd76e1ce41366a4bbe8a49b02a028
13fe9d84310e77f13a6d184dbf1232f3
e4da3b7fbbce2345d7772b0674a318d5
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

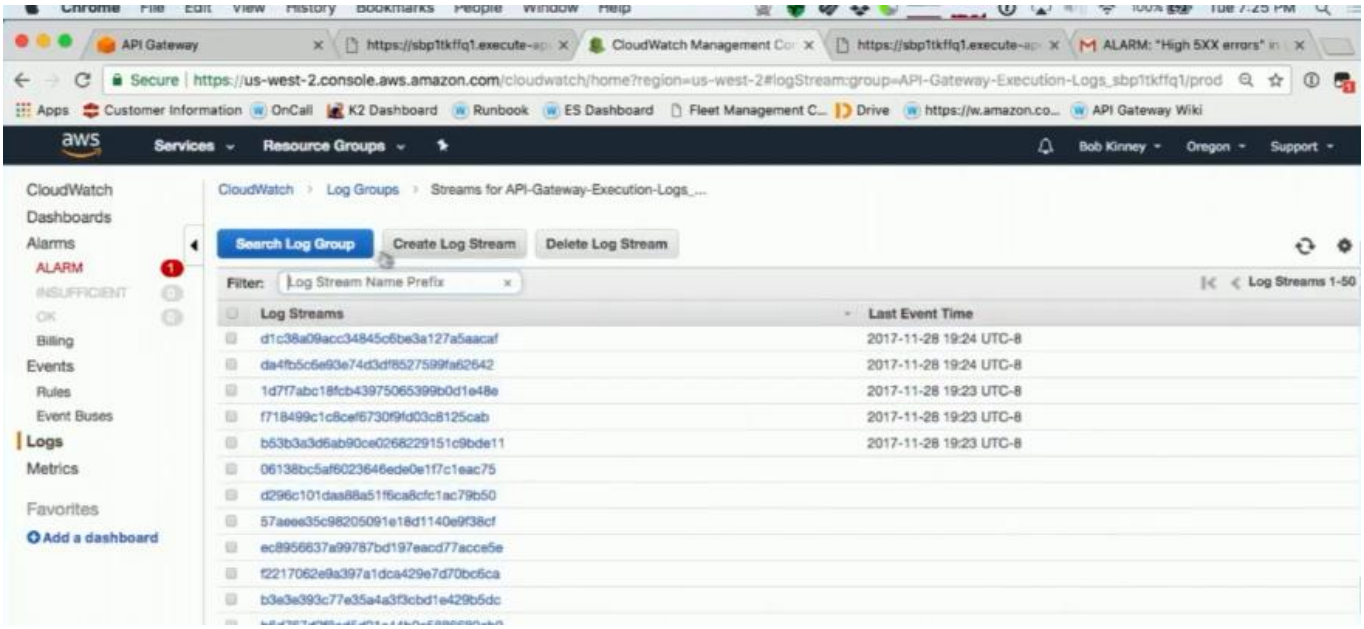

Secure | https://us-west-2.console.aws.amazon.com/cloudwatch/home?region=us-west-2#logEventViewer:group=API-Gateway-Access-Logs;stream=6f3...
Services
Resource Groups
Bob Kinney
Oregon
Support
CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
CloudWatch > Log Groups > API-Gateway-Access-Logs > 6f3ef77ac0e3619e98159e9b6febf557
Expand all
Row
Text
Filter events
all
30s
5m
1h
6h
1d
1w
custom
Time (UTC +00:00)
Message
2017-11-29
No older events found at the moment. Retry.
03:24:05
208.78.236.25 - - [29/Nov/2017:03:24:05 +0000] "GET /unauth HTTP/1.1" 500 36 c0d6971a-d4b4-11e7-8268-c7eeccffa1da
No newer events found at the moment. Retry.

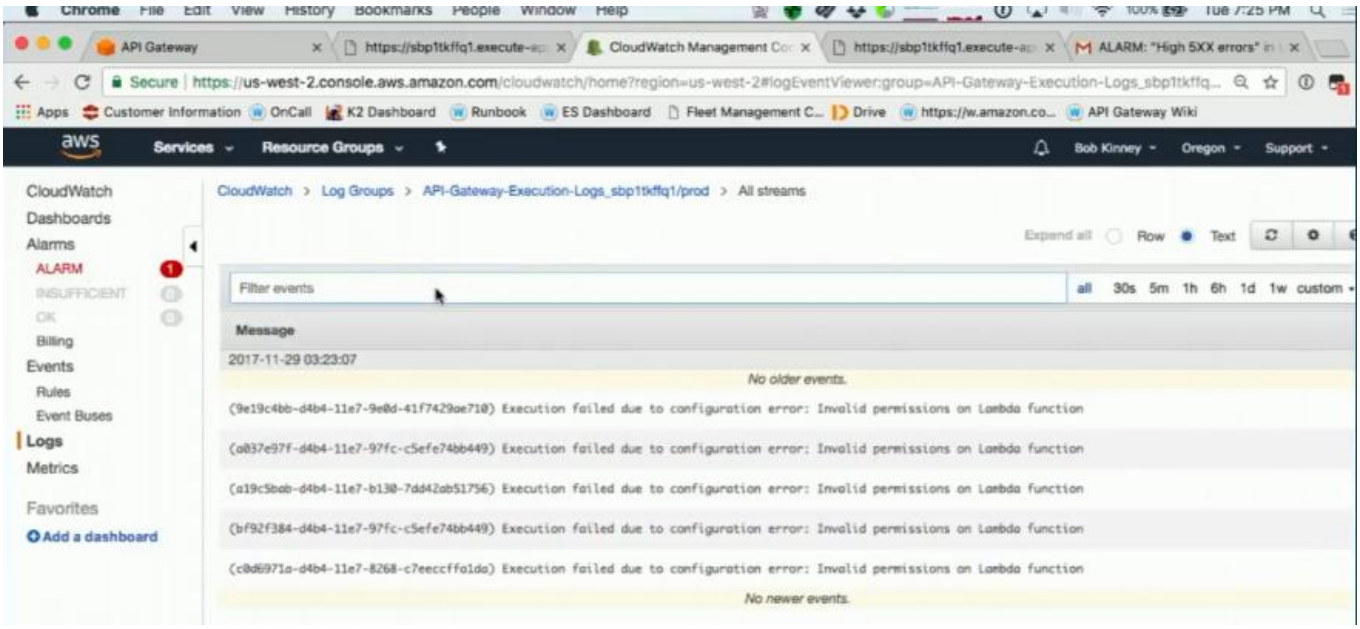


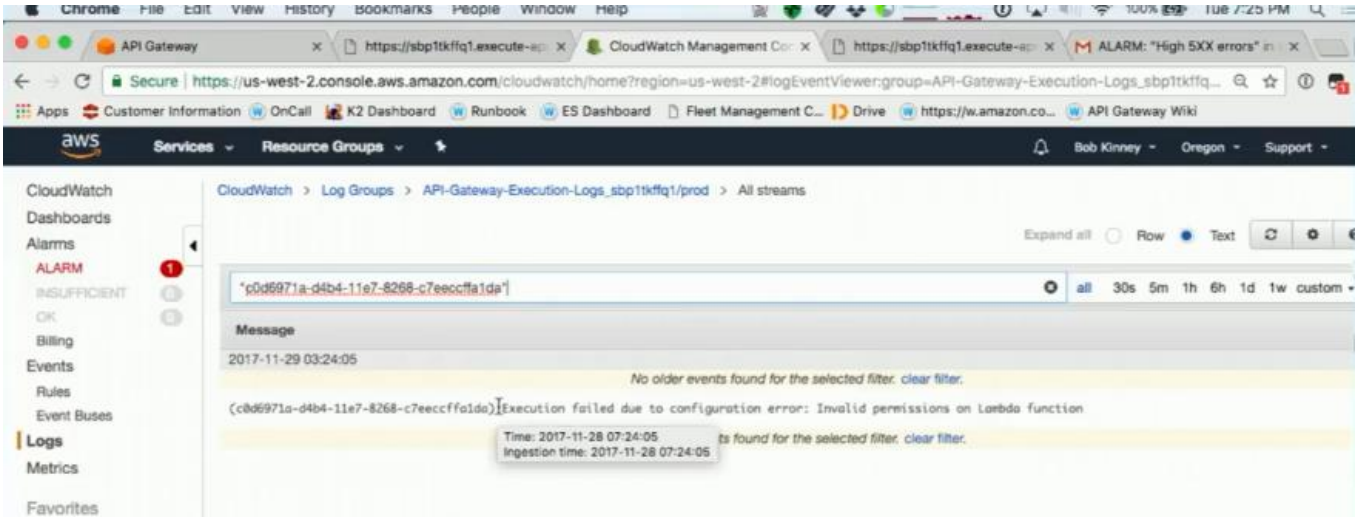
We get a single log line with all the details we need for debugging issues. This also helps us search for the actual logs as below for more context



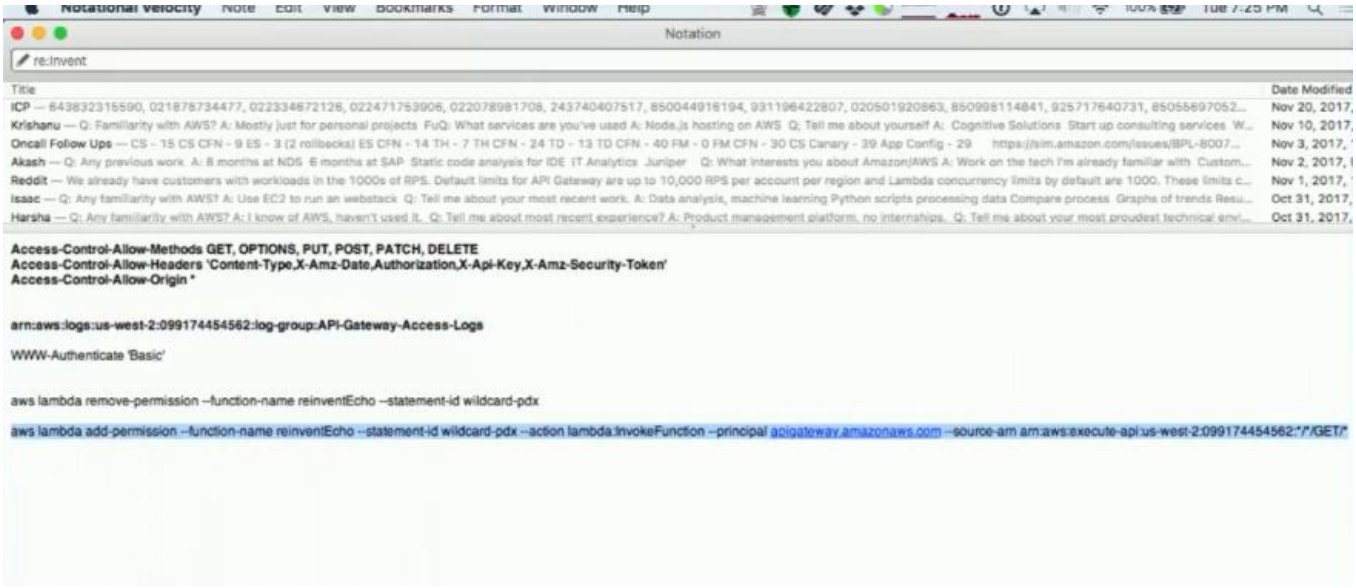
This is our regular detailed logs



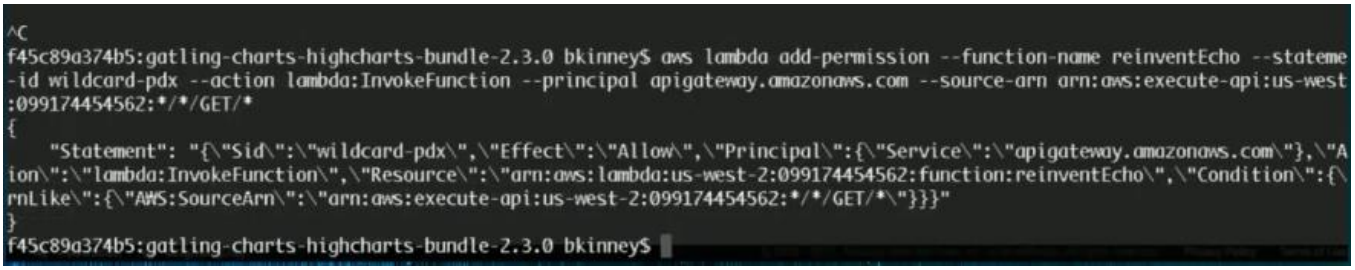




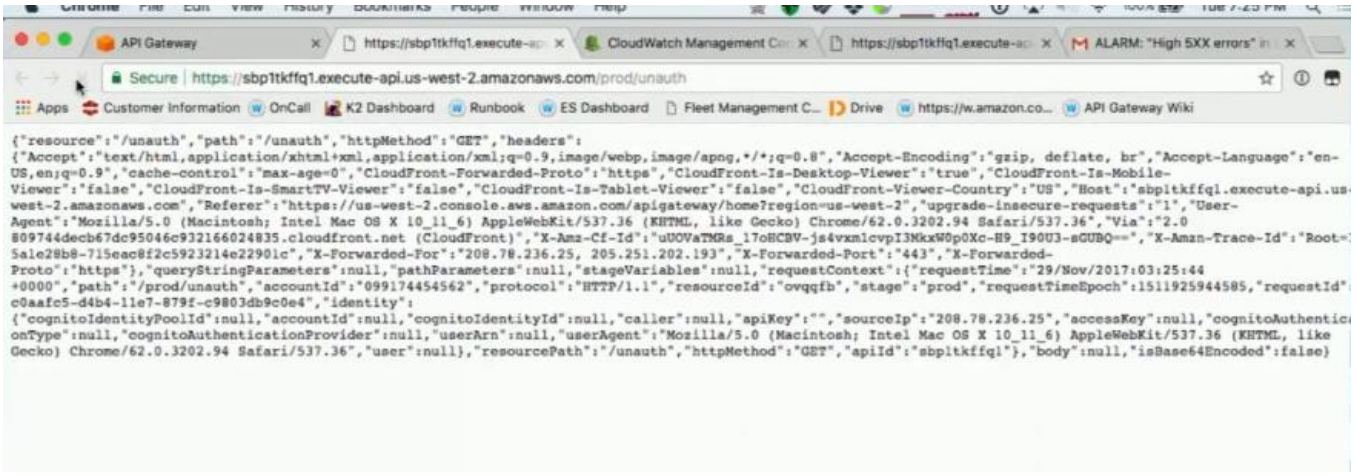
We can now see the actual error message that was recorded by API Gateway for the 500 response. We can see that the permissions we have set are not correct



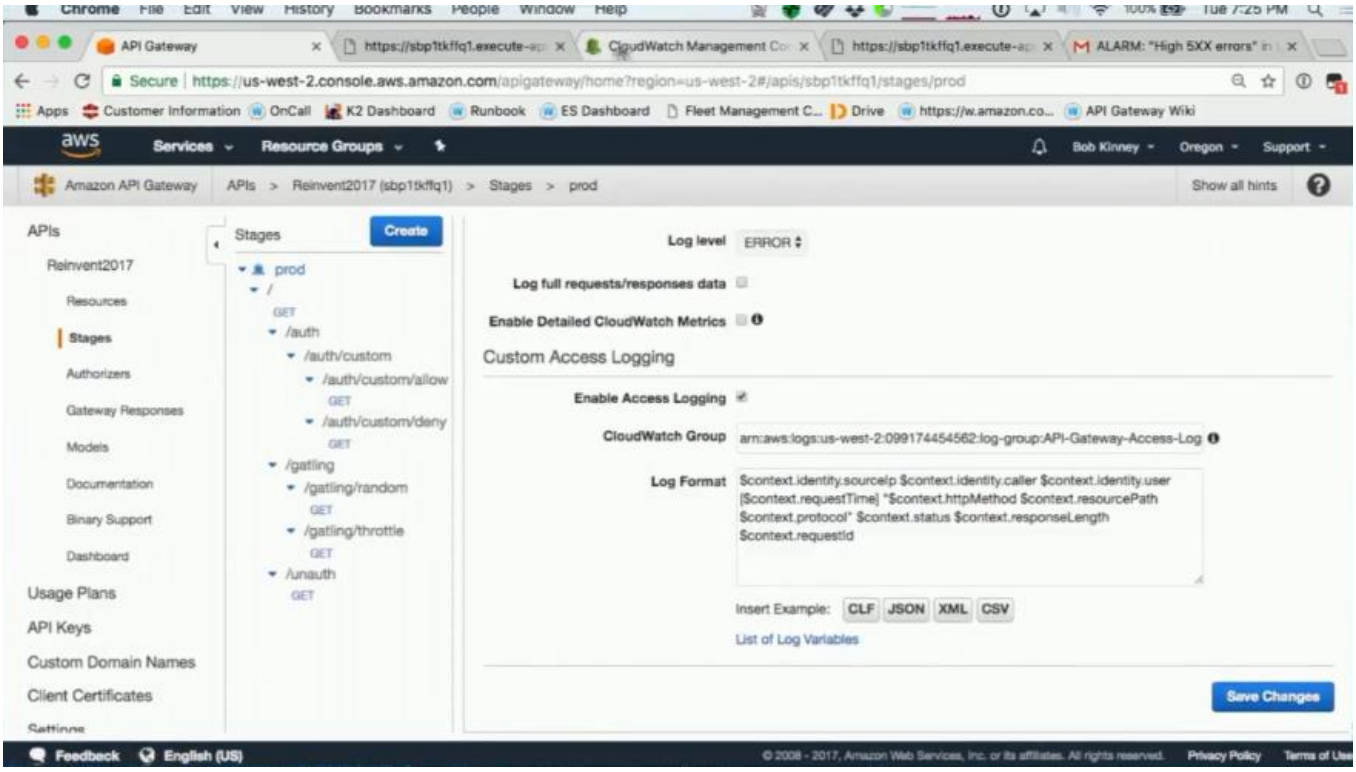
We can now set the correct permissions to resolve this problem

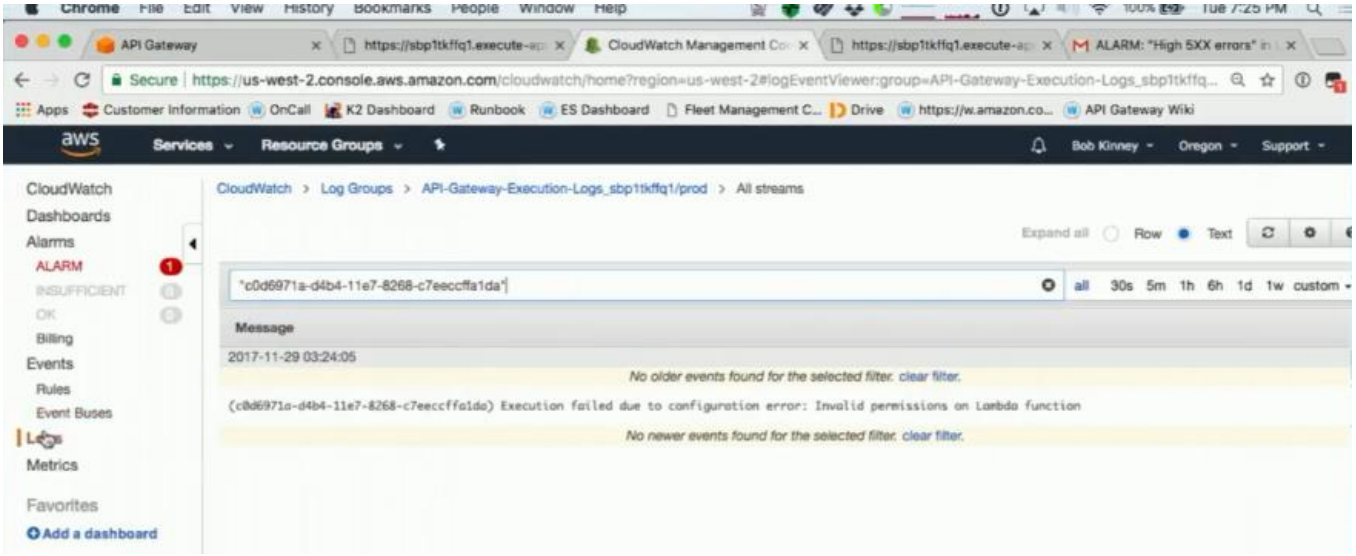


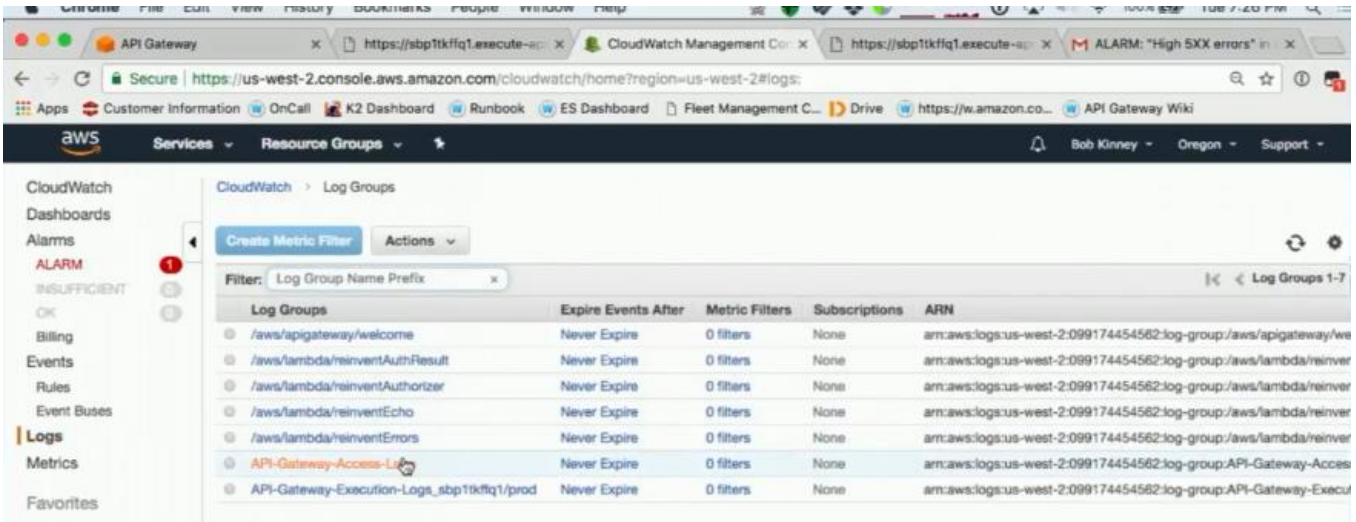
We then use the CLI to set the correct permissions

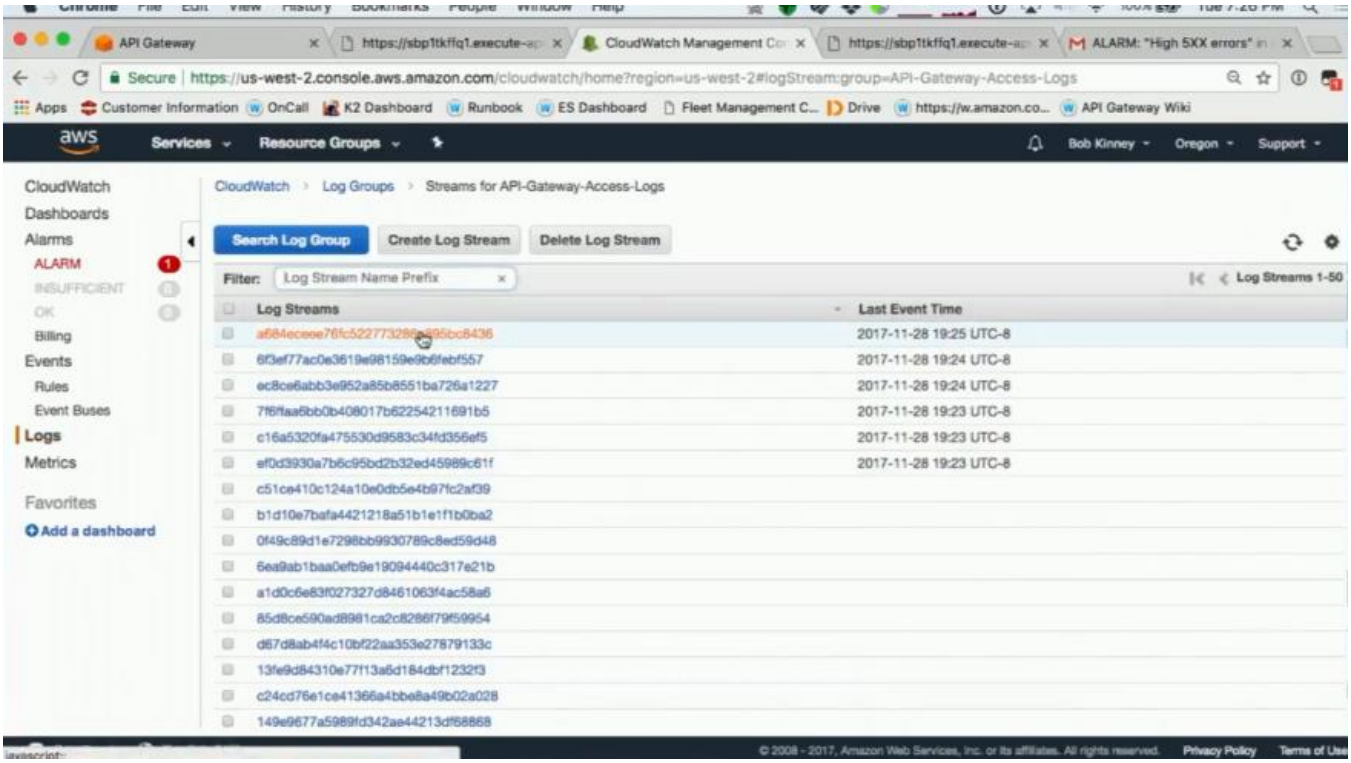


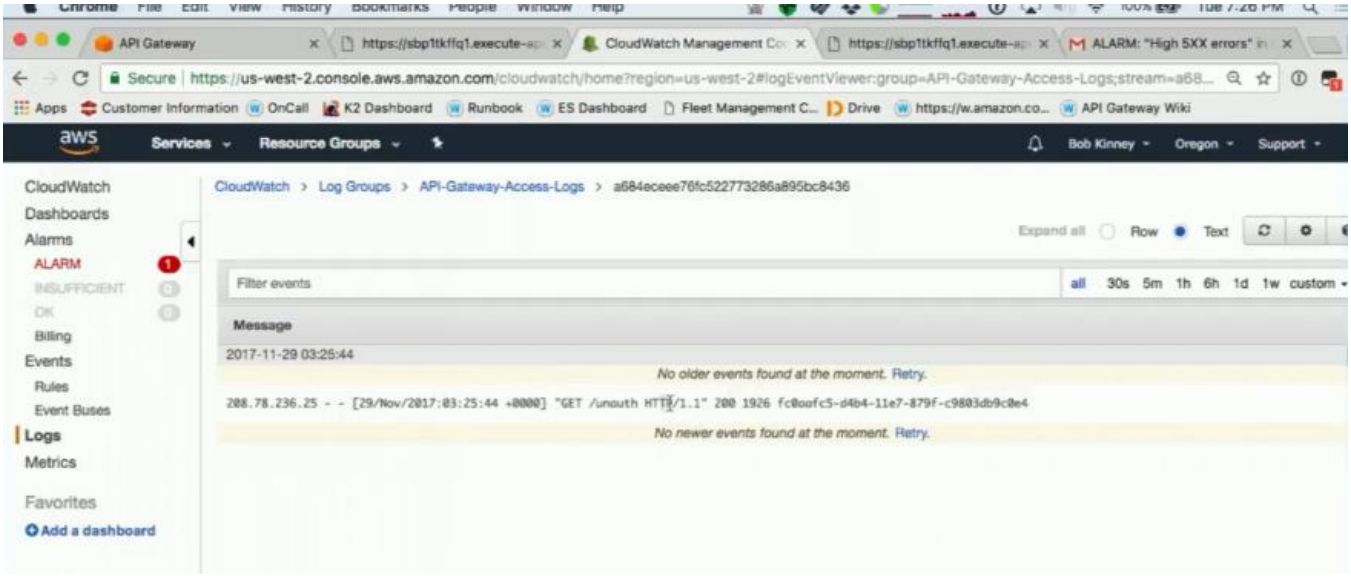
We are now getting back a 200 response with the correct data from that endpoint


Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Stages > prod
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Stages
Create
prod
/
GET
/auth
/auth/custom
/auth/custom/allow
GET
/auth/custom/deny
GET
/gatling
/gatling/random
GET
/gatling/throttle
GET
/unauth
GET
Log level
ERROR
Log full requests/responses data
Enable Detailed CloudWatch Metrics
Custom Access Logging
Enable Access Logging
CloudWatch Group
arn:aws:logs:us-west-2:099174454562:log-group:API-Gateway-Access-Log
Log Format
$context.identity.sourceIp $context.identity.caller $context.identity.user [$context.requestTime] "$context.httpMethod $context.resourcePath $context.protocol" $context.status $context.responseLength $context.requestId
Insert Example:
CLF
JSON
XML
CSV
List of Log Variables
Save Changes
Feedback
English (US)
Privacy Policy
Terms of Use


CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
CloudWatch > Log Groups > API-Gateway-Execution-Logs_sbp1tkffq1/prod > All streams
Expand all
Row
Text
"c0d6971a-d4b4-11e7-8268-c7eeccffa1da"
all
30s
5m
1h
6h
1d
1w
custom
Message
2017-11-29 03:24:05
No older events found for the selected filter. clear filter.
(c0d6971a-d4b4-11e7-8268-c7eeccffa1da) Execution failed due to configuration error: Invalid permissions on Lambda function
No newer events found for the selected filter. clear filter.

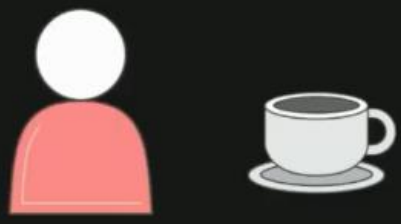

CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
CloudWatch > Log Groups
Create Metric Filter
Actions
Filter:
Log Group Name Prefix
Log Groups 1-7
Log Groups
Expire Events After
Metric Filters
Subscriptions
ARN
/aws/apigateway/welcome
/aws/lambda/reinventAuthResult
/aws/lambda/reinventAuthorizer
/aws/lambda/reinventEcho
/aws/lambda/reinventErrors
API-Gateway-Access-Log
API-Gateway-Execution-Logs_sbp1tkffq1/prod
Never Expire
0 filters
None



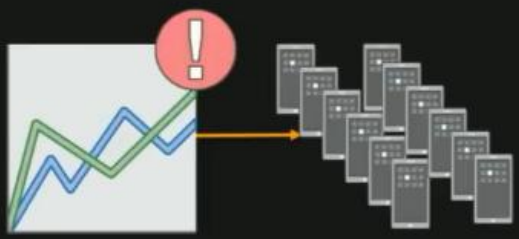


We can now see a 200 response with the same data and request Ids. This is a great way to look at the errors being generated and then use the request ids to see what is happening using the logs in the API Gateway or lambdas.

Check in with Doug



Thanks to logging, Doug now knows that his API is generating errors during peak loads because there's spurious traffic hitting a legacy API method at a much higher than expected rate due to a bug in a previous version of the app.



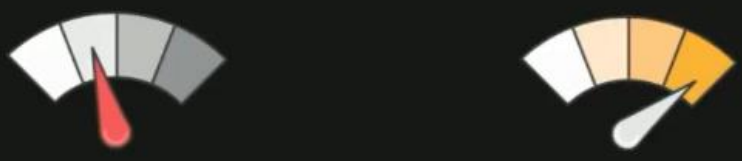
He now needs a way to to limit the traffic from those devices to let other traffic through from updated devices.

Protecting Your API: Throttling

API Gateway throttling

Three levels of throttling for APIs

- API Key level throttling—configurable in usage plan
- Method level throttling—configurable in stage settings
- Account level throttling—limits can be increased



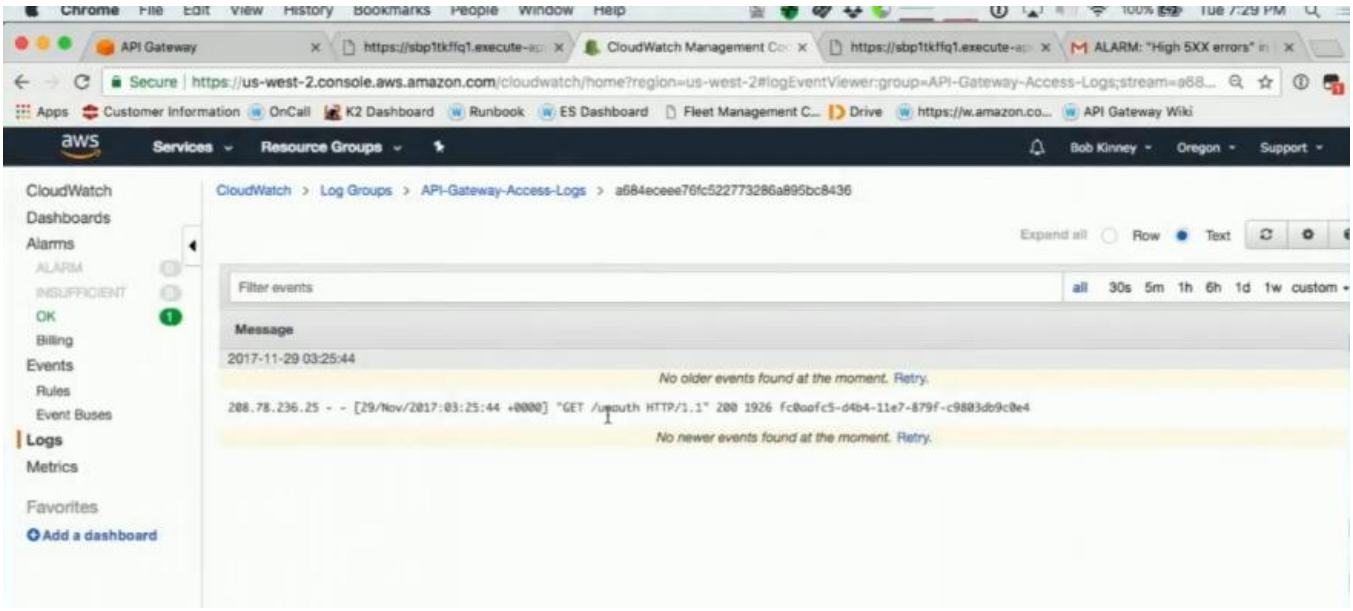
API Gateway throttling

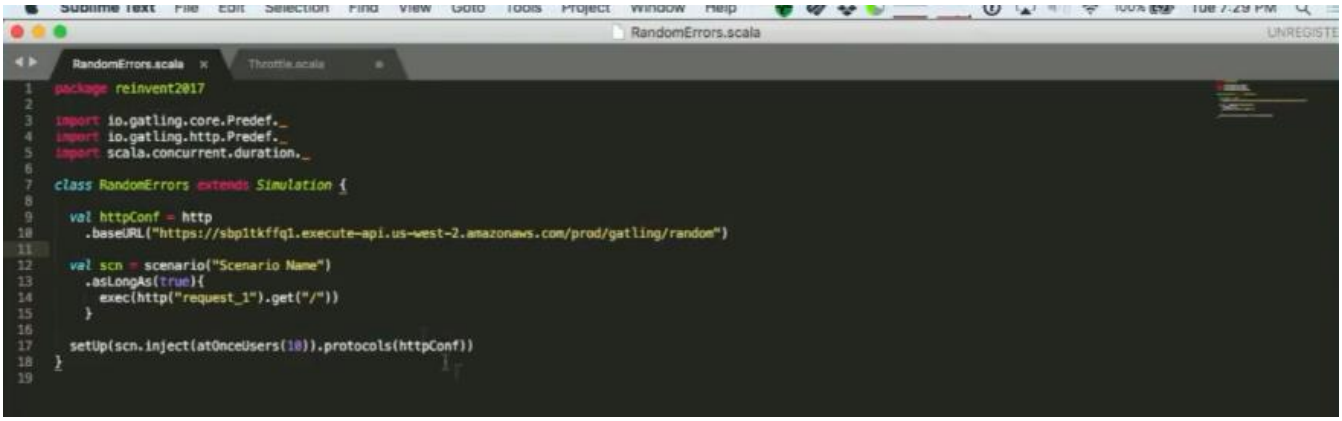
Token bucket algorithm

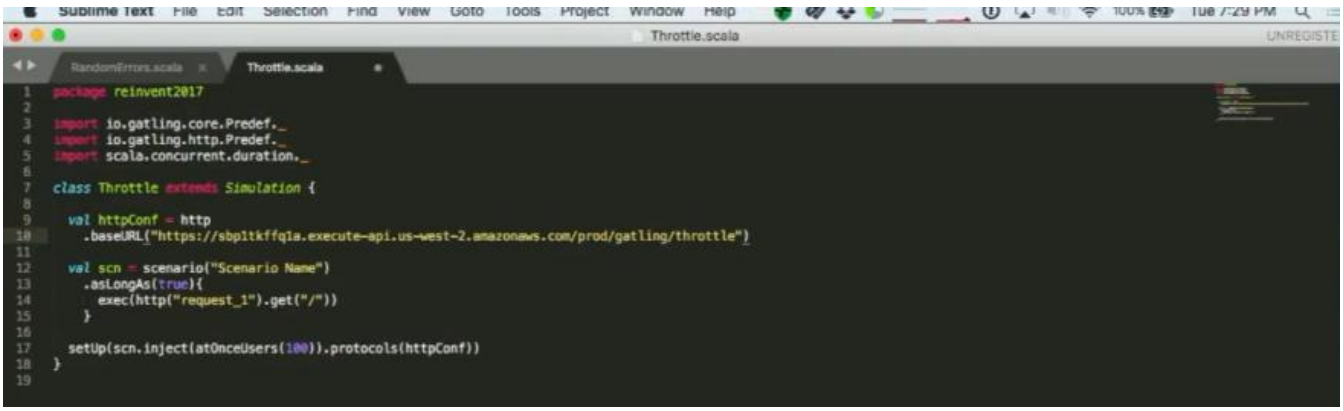
- Burst—the maximum size of the bucket
- Rate—the number of tokens added to the bucket

Demo

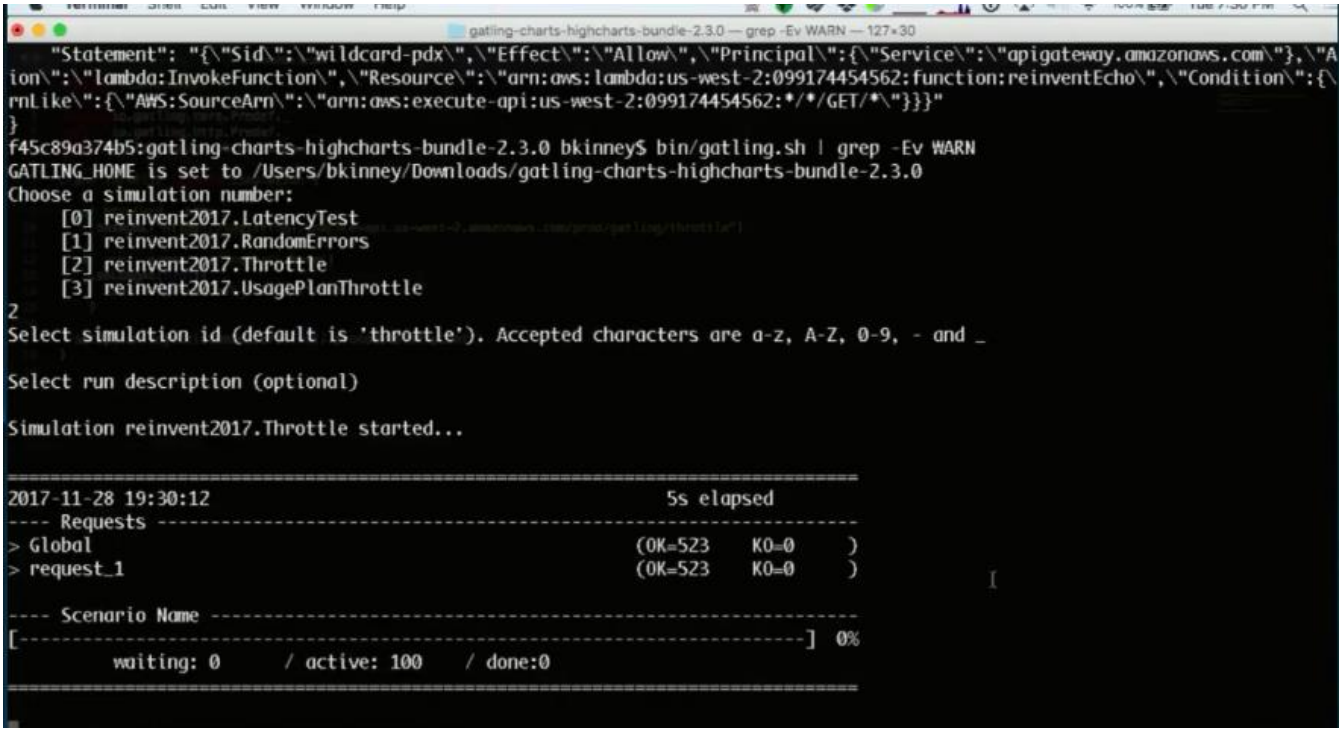
Let us now see a demo about enabling throttling

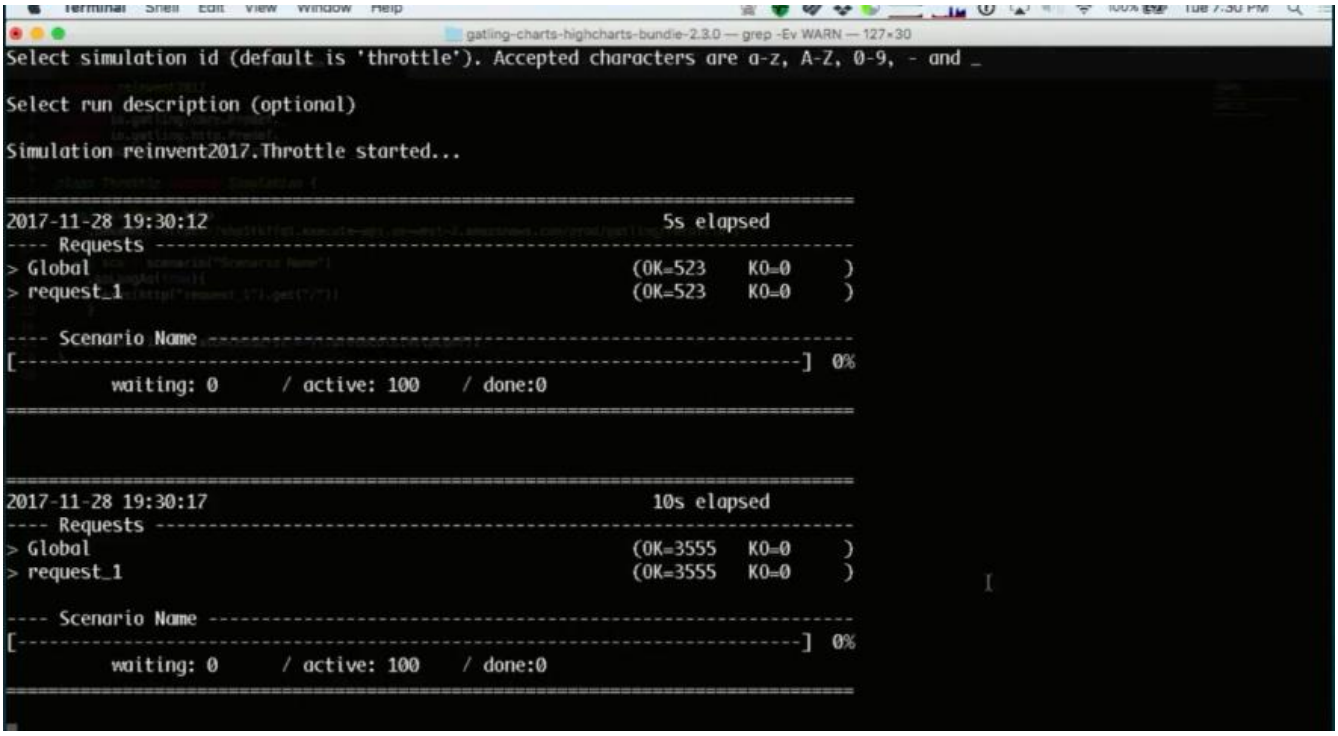




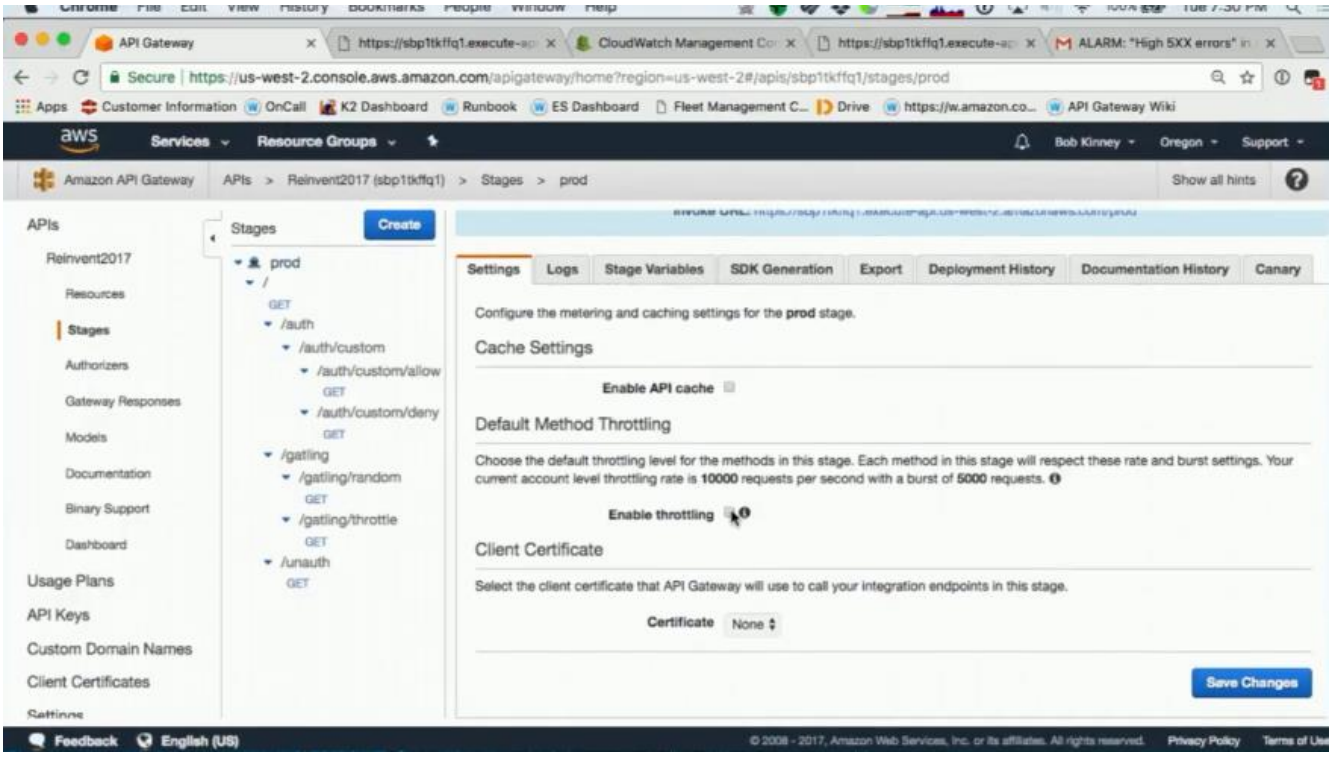


We have another Gatling script that hits a specific method in our API, we are going to start this up and then enable throttling

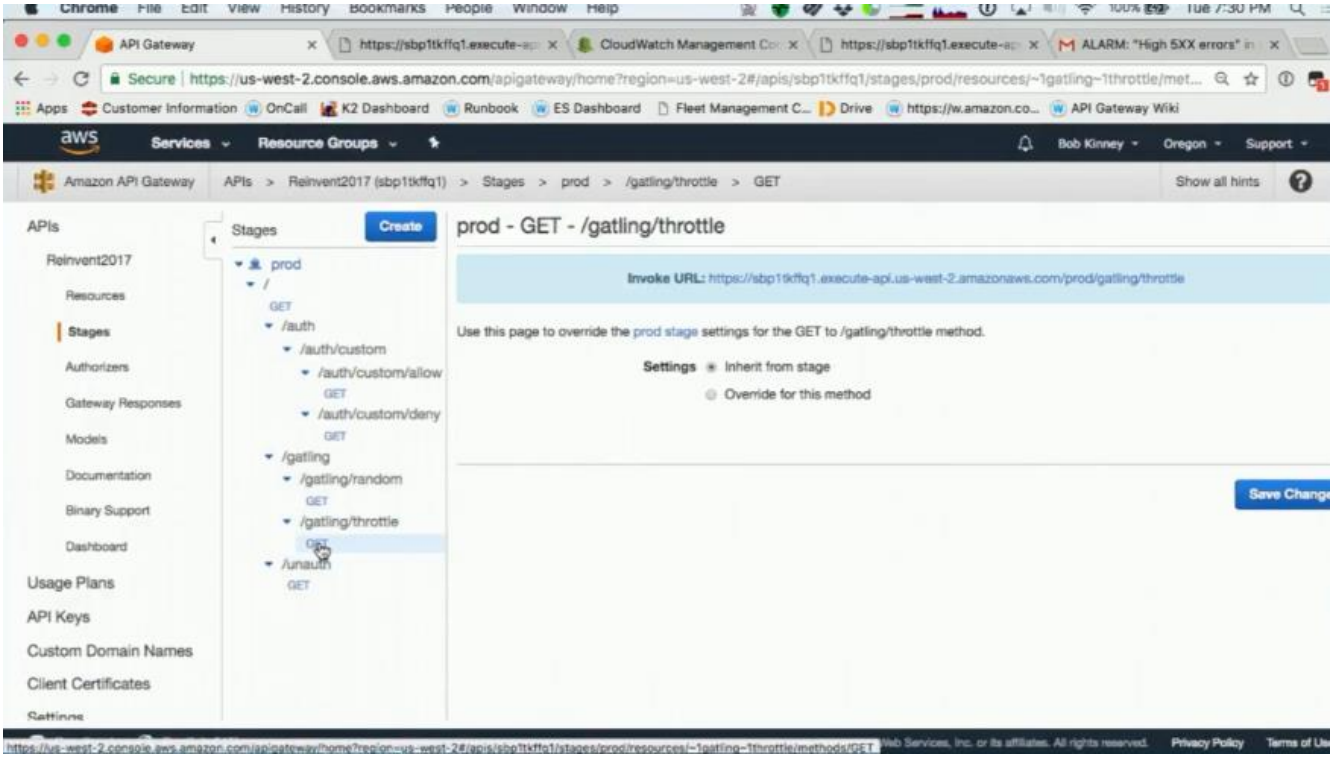


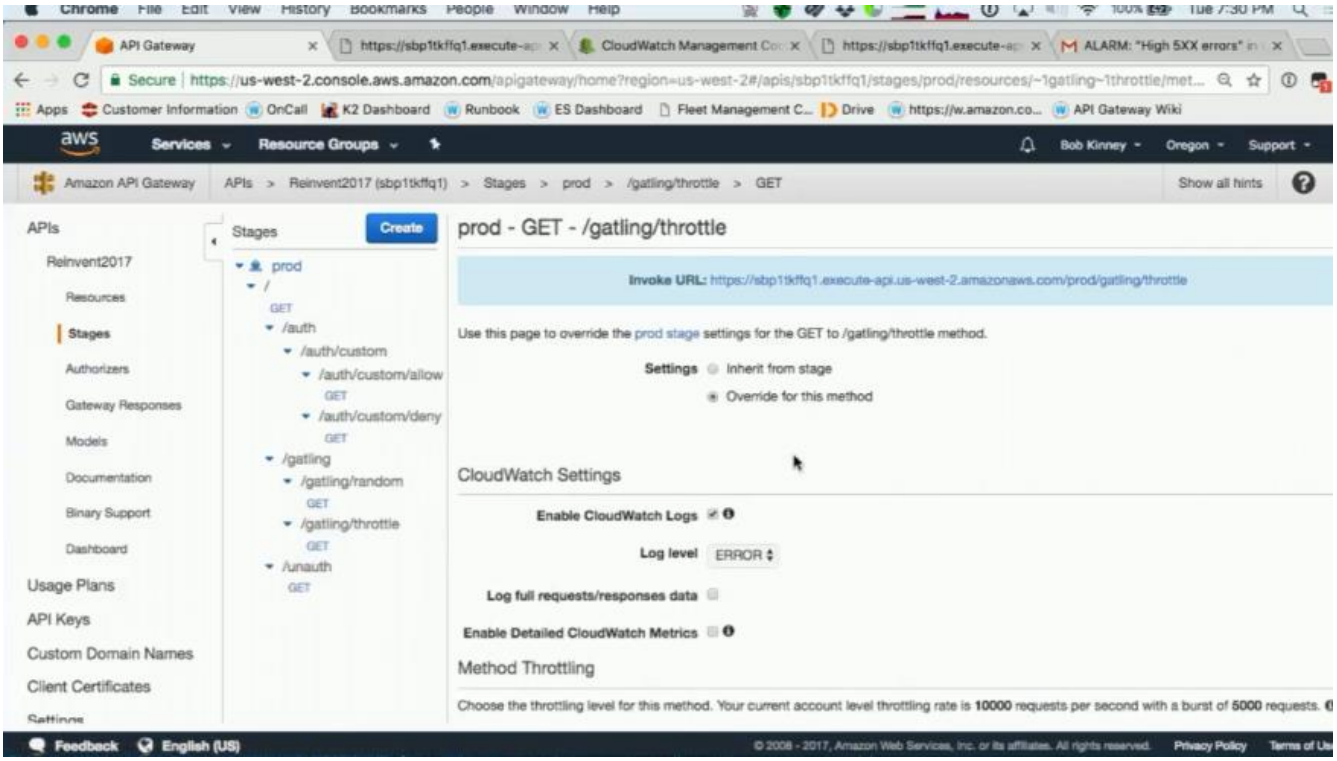


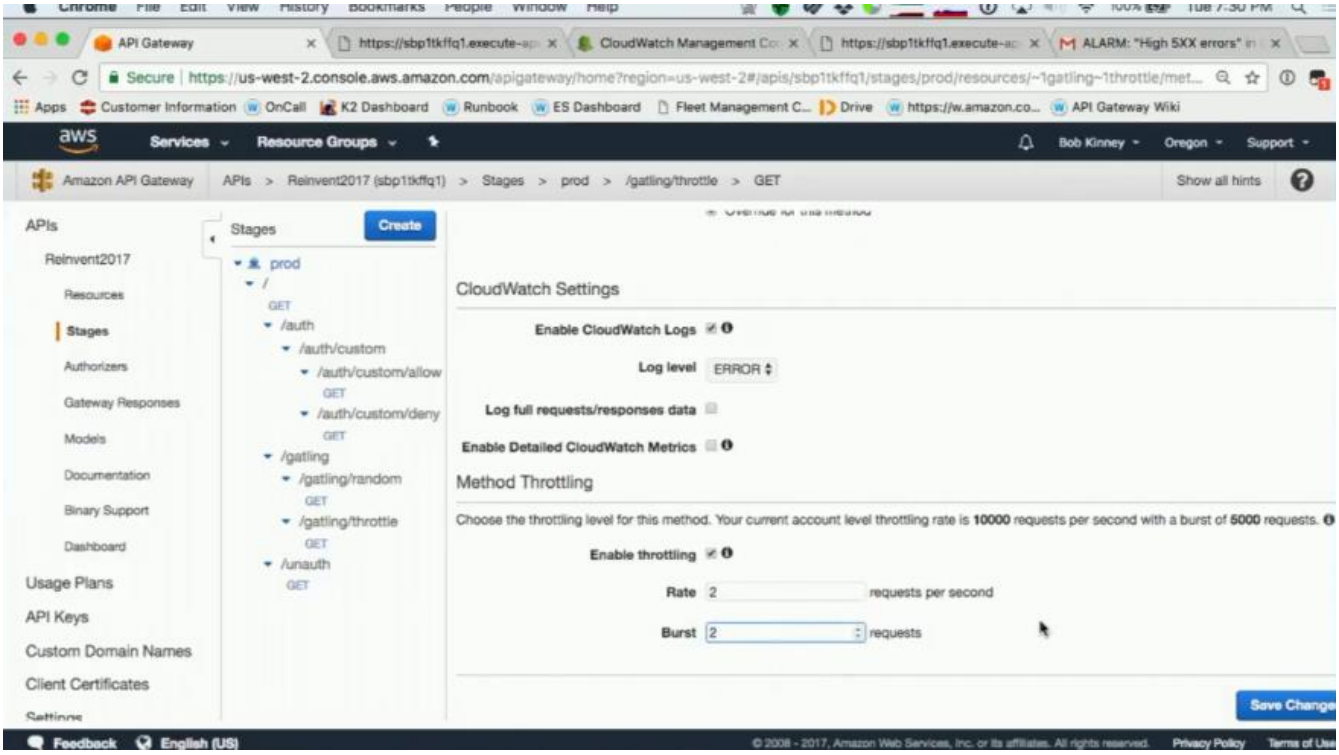
All of these requests are generating 200 status codes for about 100 requests per second RPS to the API endpoint, this is because we have not enabled any request throttling at all. We can apply throttling to keep the RPS lower allowed to this endpoint.



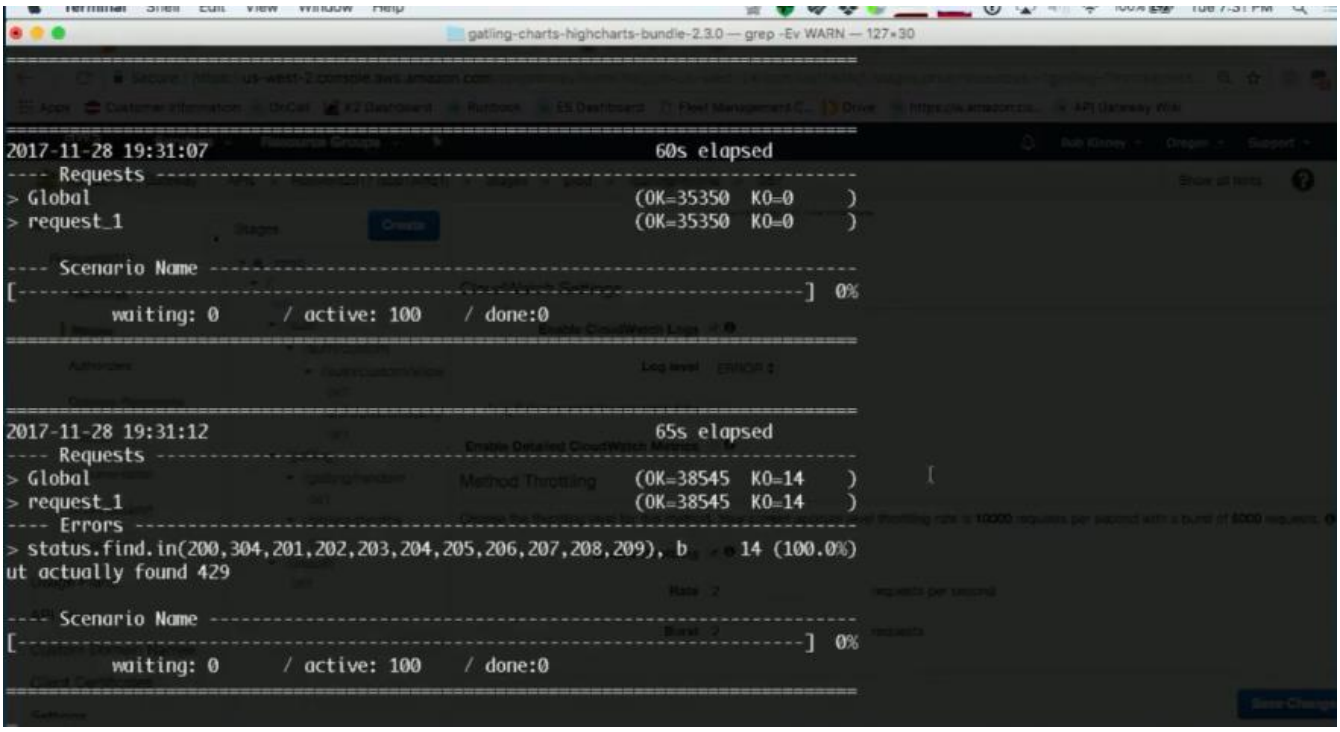
We can set throttling on the entire stage API as above, or set throttling on the specific GET endpoint only as below


Chrome File Edit View History Bookmarks People Window Help
API Gateway
CloudWatch Management Co
ALARM: "High 5XX errors"
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/stages/prod/resources/~1gatling~1throttle/met...
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
Amazon API Gateway APIs > Reinvent2017 (sbp1tkffq1) > Stages > prod > /gatling/throttle > GET Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Stages
Create
prod
/
GET
/auth
/auth/custom
/auth/custom/allow
GET
/auth/custom/deny
GET
/gatling
/gatling/random
GET
/gatling/throttle
GET
/unauth
GET
prod - GET - /gatling/throttle
Invoke URL: https://sbp1tkffq1.execute-api.us-west-2.amazonaws.com/prod/gatling/throttle
Use this page to override the prod stage settings for the GET to /gatling/throttle method.
Settings
Inherit from stage
Override for this method
Save Change
https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/stages/prod/resources/~1gatling~1throttle/methods/GET
Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

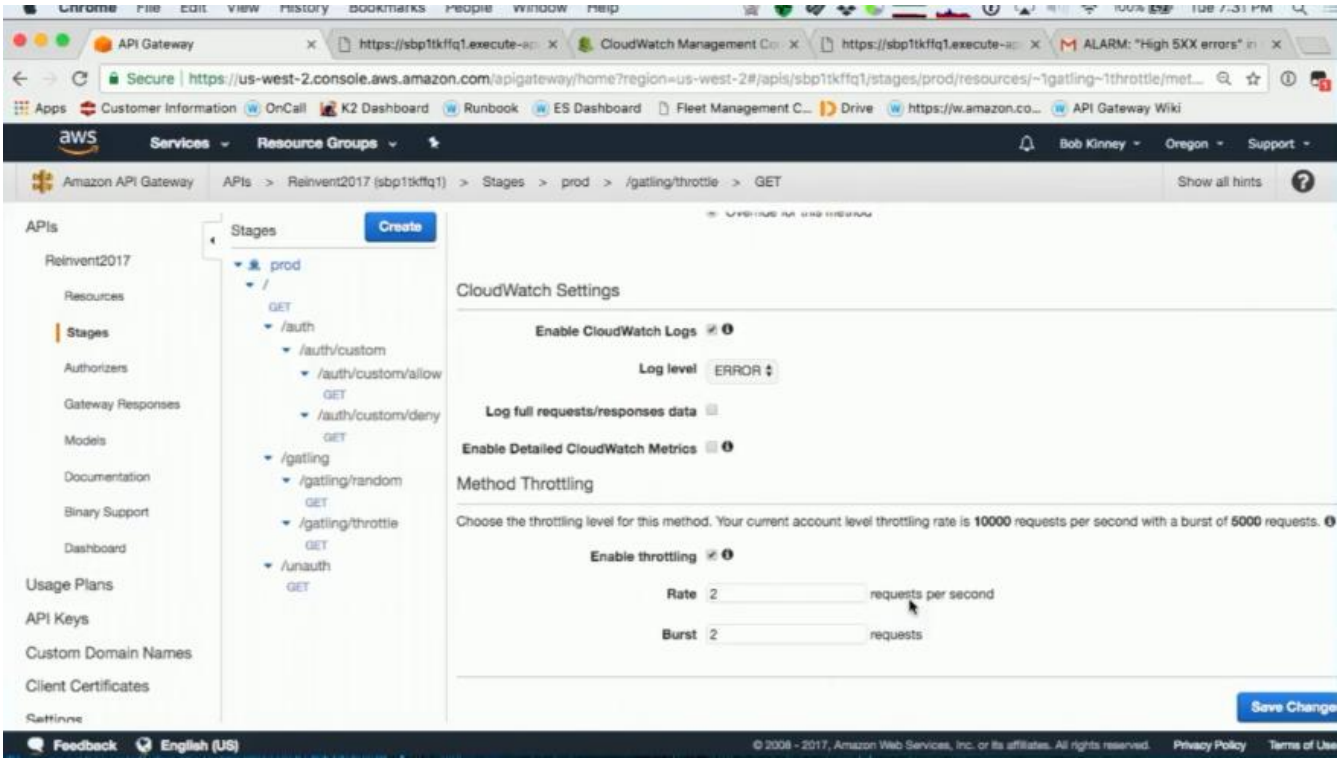

Chrome File Edit View History Bookmarks People Window Help
API Gateway
CloudWatch Management Co
ALARM: "High 5XX errors"
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/stages/prod/resources/~1gatling~1throttle/met...
Apps Customer Information OnCall K2 Dashboard Runbook ES Dashboard Fleet Management C... Drive https://w.amazon.co... API Gateway Wiki
aws Services Resource Groups Bob Kinney Oregon Support
Amazon API Gateway APIs > Reinvent2017 (sbp1tkffq1) > Stages > prod > /gatling/throttle > GET Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Stages
Create
prod
/
GET
/auth
/auth/custom
/auth/custom/allow
GET
/auth/custom/deny
GET
/gatling
/gatling/random
GET
/gatling/throttle
GET
/unauth
GET
prod - GET - /gatling/throttle
Invoke URL: https://sbp1tkffq1.execute-api.us-west-2.amazonaws.com/prod/gatling/throttle
Use this page to override the prod stage settings for the GET to /gatling/throttle method.
Settings
Inherit from stage
Override for this method
CloudWatch Settings
Enable CloudWatch Logs
Log level
ERROR
Log full requests/responses data
Enable Detailed CloudWatch Metrics
Method Throttling
Choose the throttling level for this method. Your current account level throttling rate is 10000 requests per second with a burst of 5000 requests.
Feedback
English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

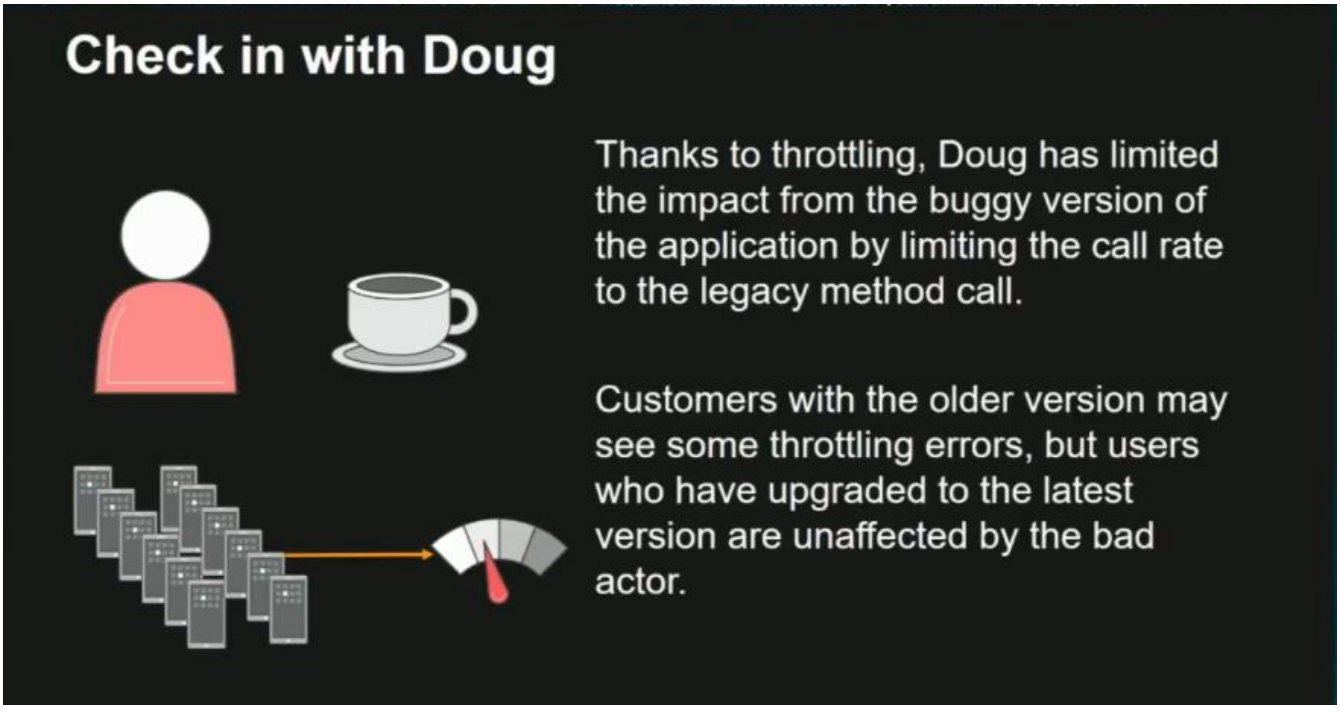


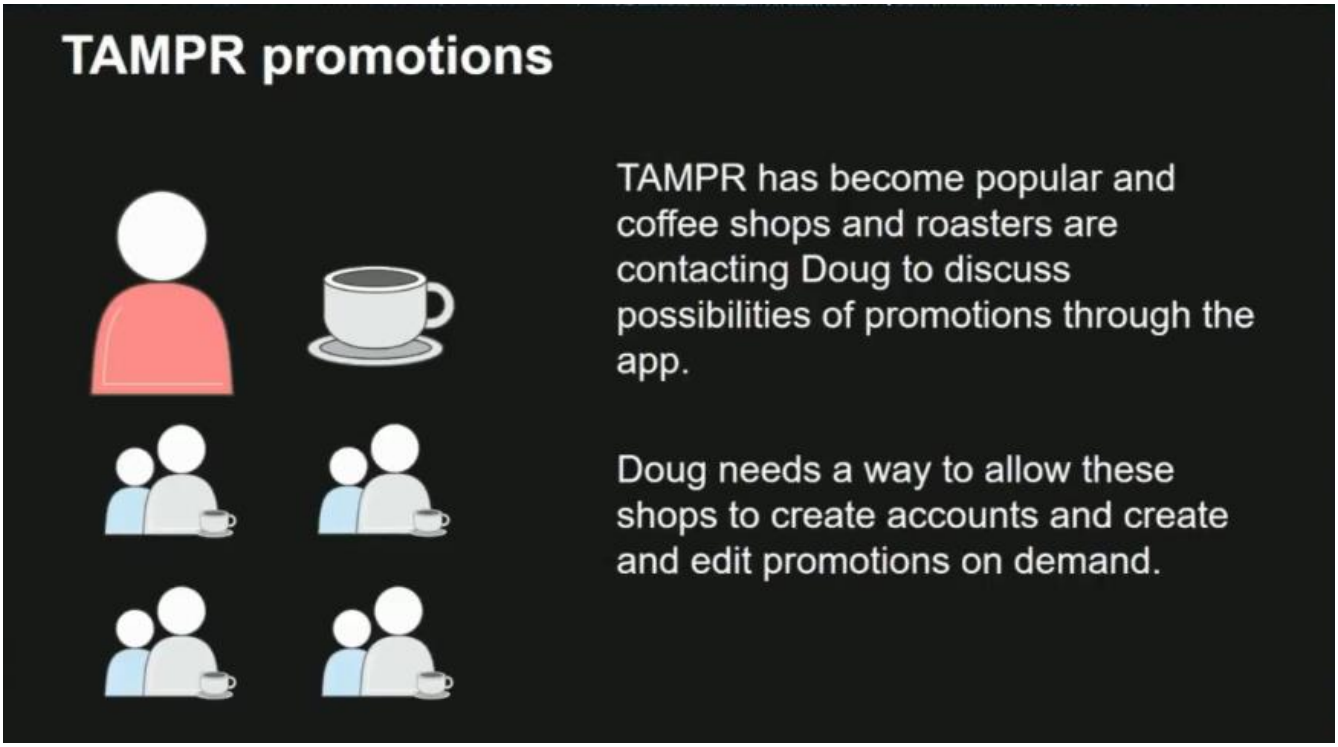
Set this to a lower RPS like 2 for Rate and 2 for Burst as above



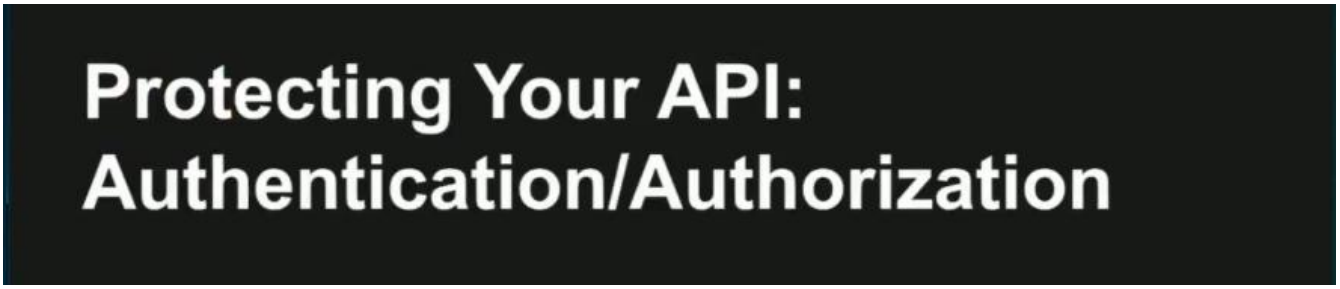
We are now getting 429 responses from the API for some requests, this will save us API calls cost.


API Gateway
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/stages/prod/resources/~1gatling~1throttle/met...
Apps
Customer Information
OnCall
K2 Dashboard
Runbook
ES Dashboard
Fleet Management C...
Drive
https://w.amazon.co...
API Gateway Wiki
aws
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Stages > prod > /gatling/throttle > GET
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Stages
Create
prod
/
GET
/auth
/auth/custom
/auth/custom/allow
GET
/auth/custom/deny
GET
/gatling
/gatling/random
GET
/gatling/throttle
GET
/unauth
GET
CloudWatch Settings
Enable CloudWatch Logs
Log level
ERROR
Log full requests/responses data
Enable Detailed CloudWatch Metrics
Method Throttling
Choose the throttling level for this method. Your current account level throttling rate is 10000 requests per second with a burst of 5000 requests.
Enable throttling
Rate
2
requests per second
Burst
2
requests
Save Changes
Feedback
English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use


Check in with Doug
Thanks to throttling, Doug has limited the impact from the buggy version of the application by limiting the call rate to the legacy method call.
Customers with the older version may see some throttling errors, but users who have upgraded to the latest version are unaffected by the bad actor.



Doug needs to open up a portion of his API so that the coffee shops wanting to run promotions can submit their promotions without him having to be a go-between.

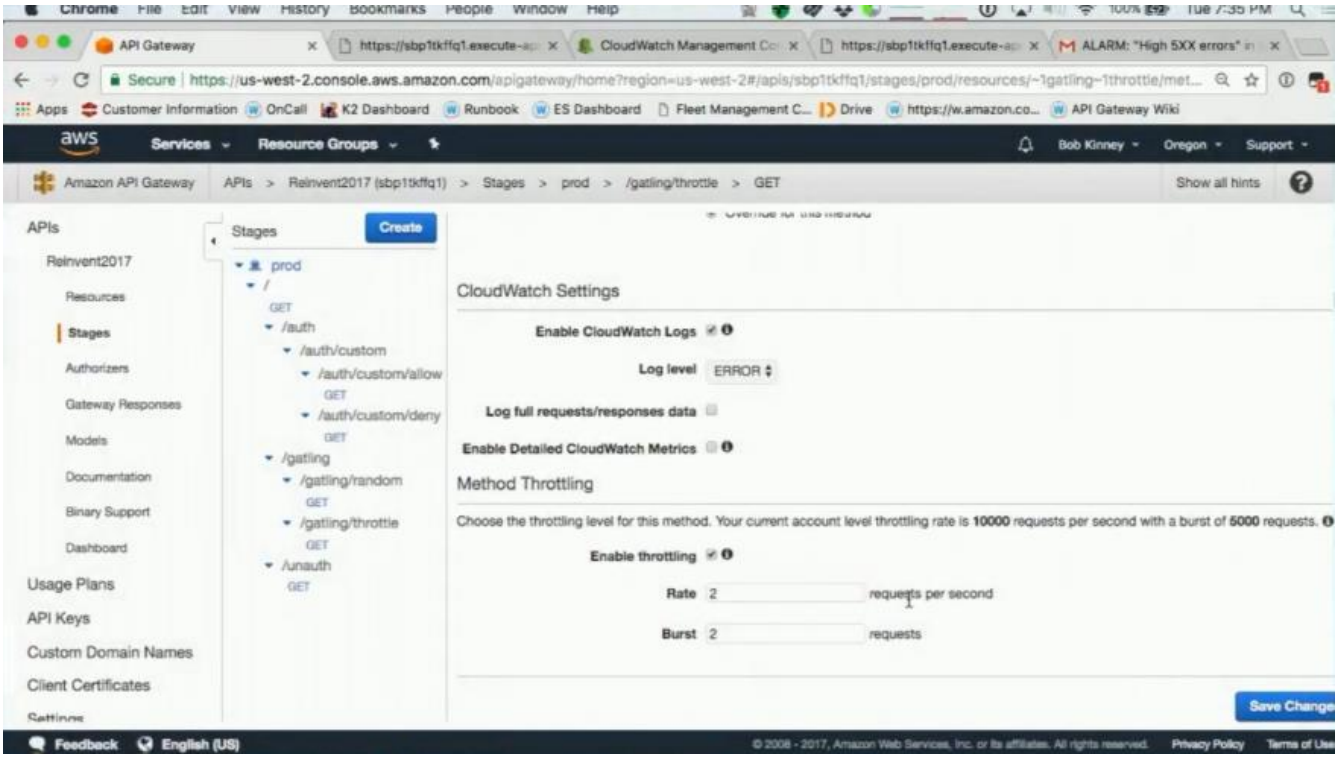


Authentication type comparison ★ NEW ★

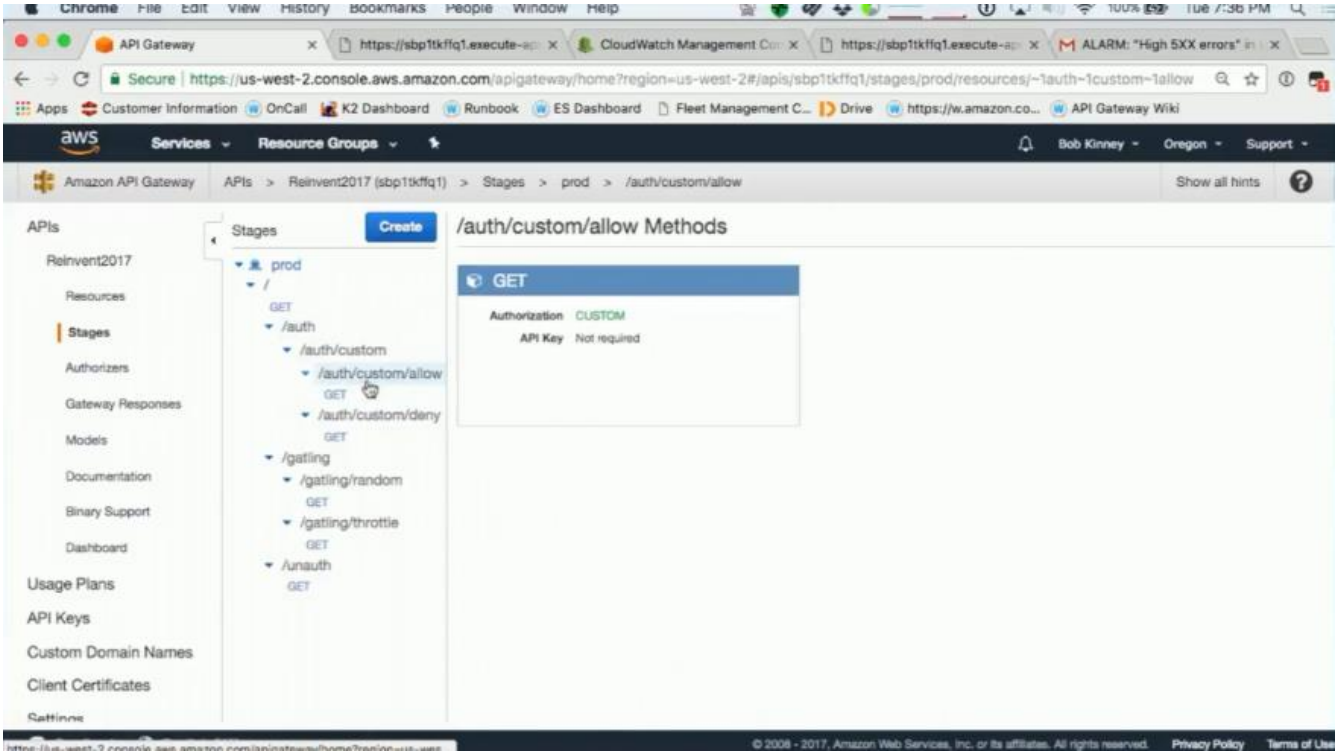
Feature	AWS_IAM	TOKEN	REQUEST	COGNITO
Authentication	X	X	X	X
Authorization	X	X	X	
Signature V4	X			
Cognito User Pools		X	X	X
Third-Party Authentication		X	X	
Multiple Header Support			X	
Additional Costs	NONE	Pay per authorizer invoke	Pay per authorizer invoke	NONE

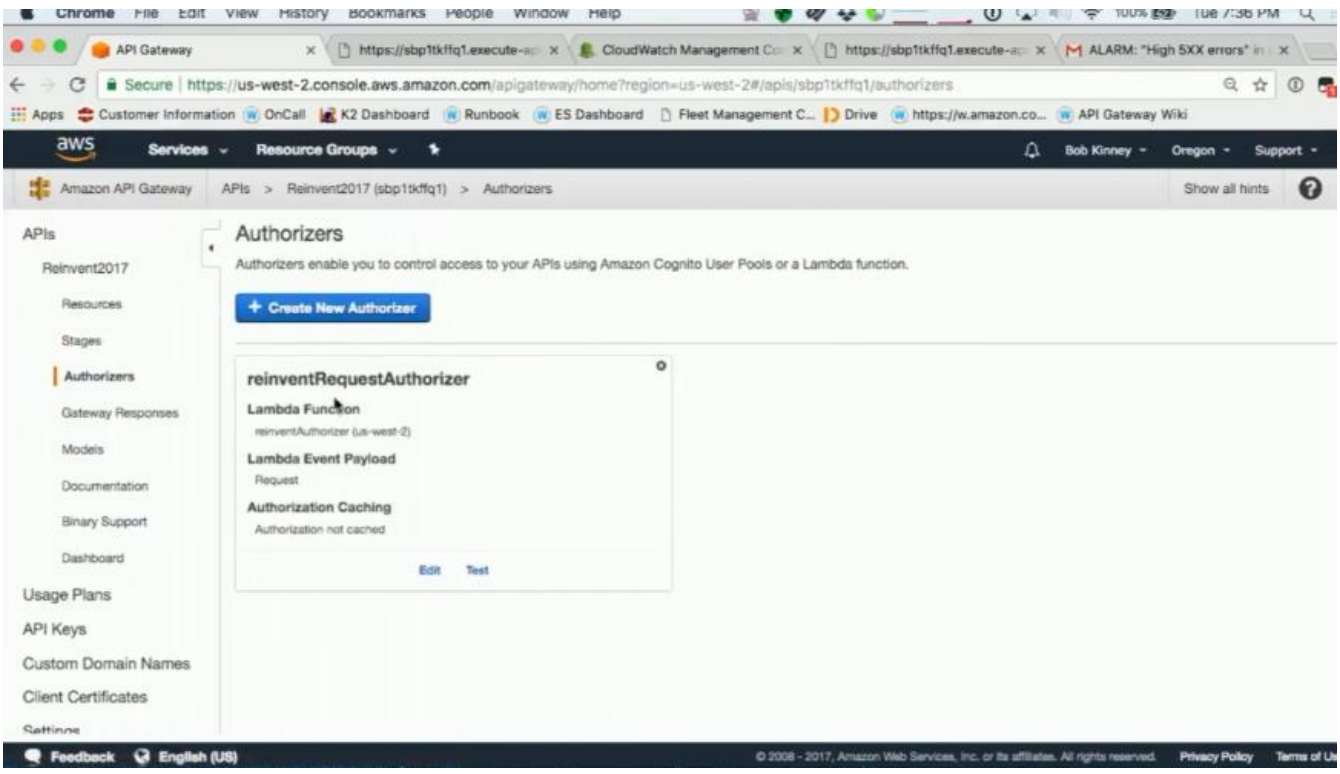
Demo

Let us now see a custom authorizer

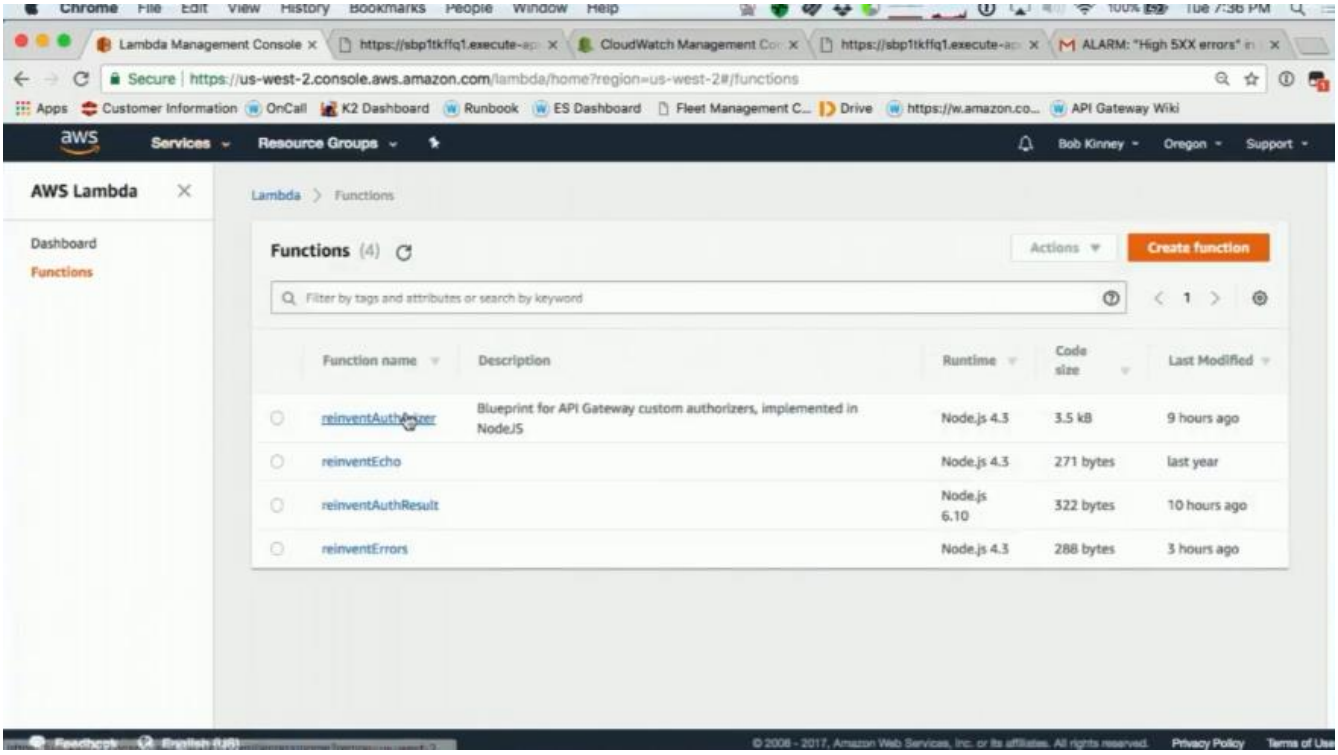


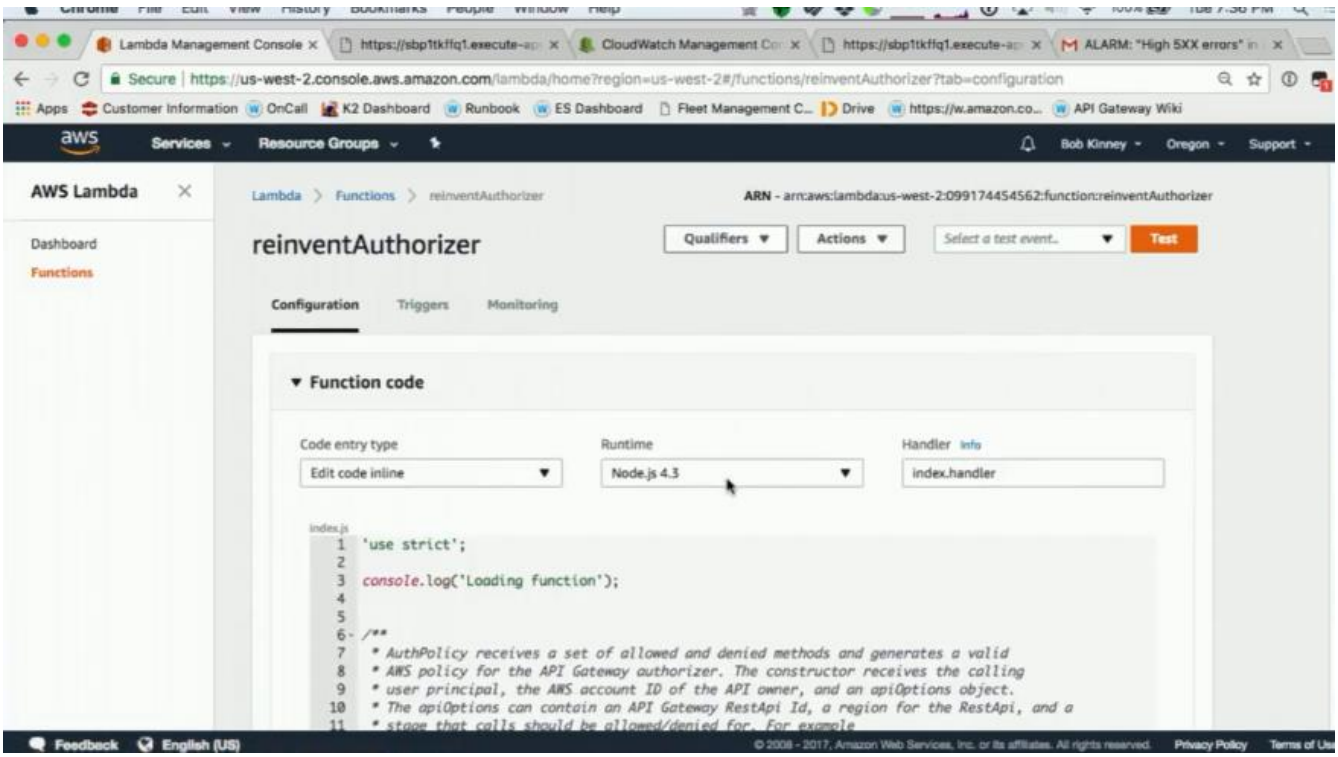
We are going to be implementing basic HTTP ***auth*** for our API Gateway API

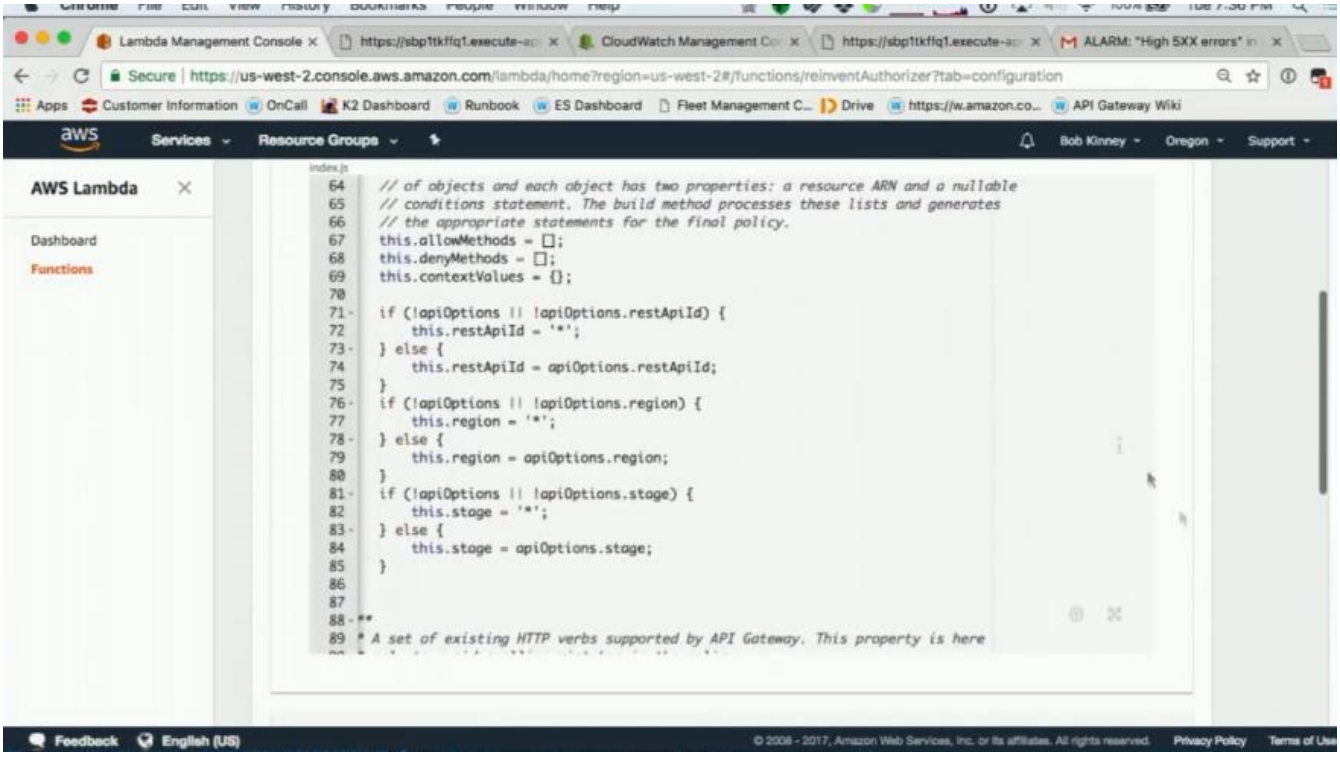


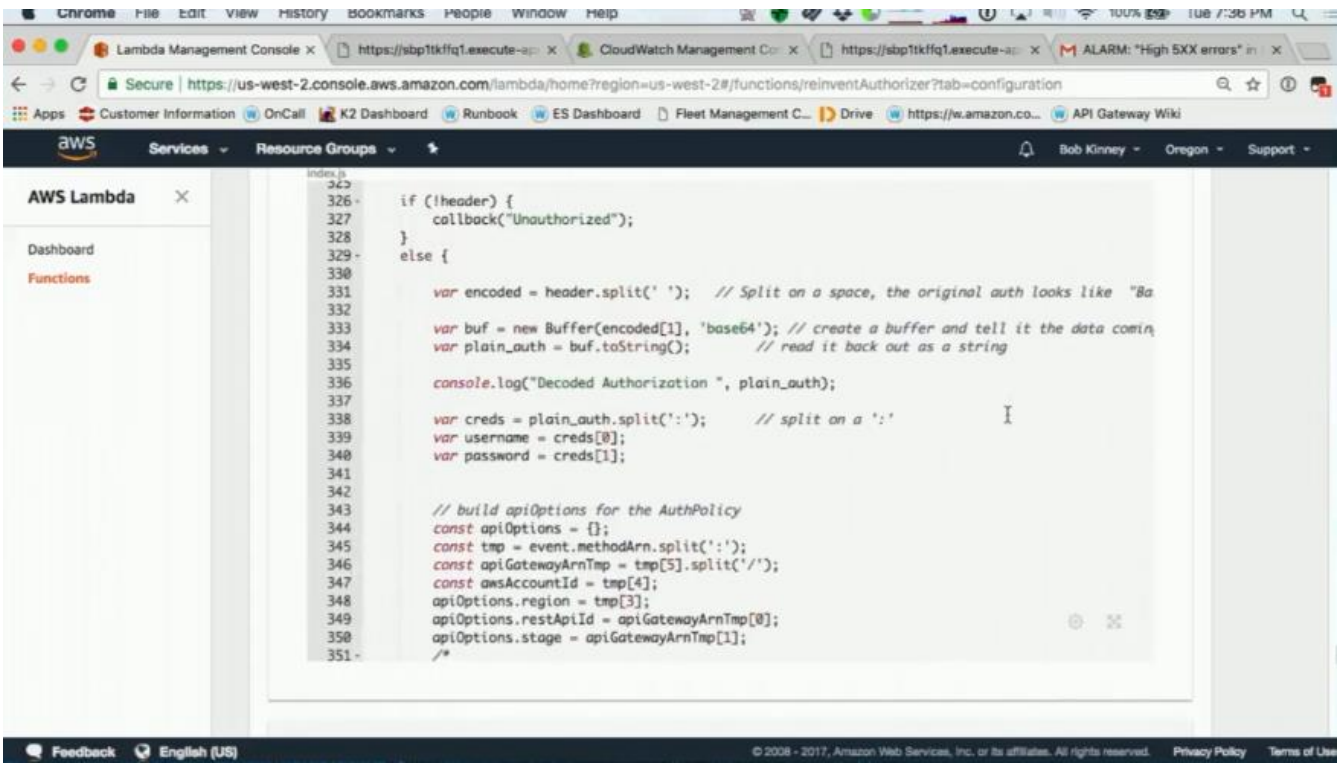


We have set up our authorizer as a Request Authorizer, basic auth does not usually require multiple headers. In this case, we are using the authorization header.

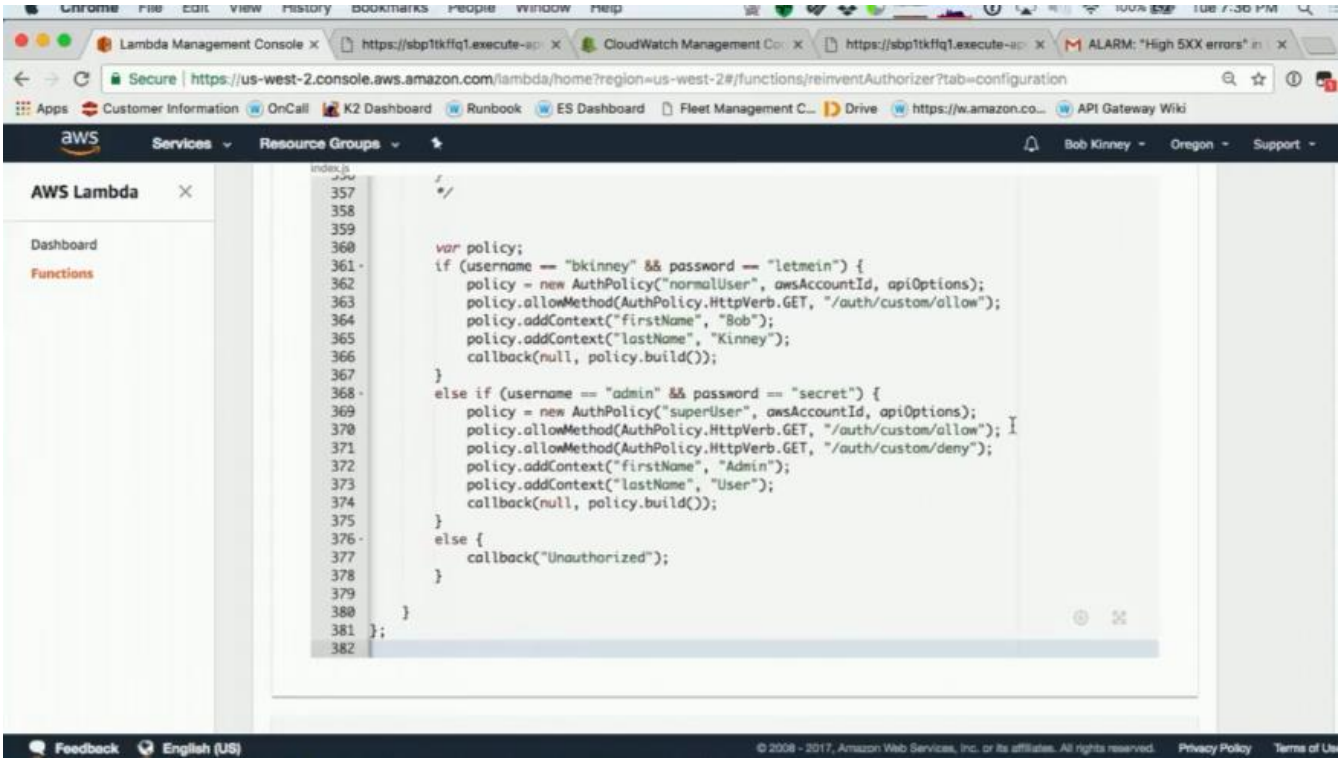

Lambda Management Console
CloudWatch Management Con
ALARM: "High 5XX errors"
Secure | https://us-west-2.console.aws.amazon.com/lambda/home?region=us-west-2#/functions
Apps
Customer Information
OnCall
K2 Dashboard
Runbook
ES Dashboard
Fleet Management C...
Drive
API Gateway Wiki
Services
Resource Groups
Bob Kinney
Oregon
Support
AWS Lambda
Dashboard
Functions
Lambda > Functions
Functions (4)
Actions
Create function
Filter by tags and attributes or search by keyword
Function name
Description
Runtime
Code size
Last Modified
reinventAuthorizer
Blueprint for API Gateway custom authorizers, implemented in Node.JS
Node.js 4.3
3.5 kB
9 hours ago
reinventEcho
Node.js 4.3
271 bytes
last year
reinventAuthResult
Node.js 6.10
322 bytes
10 hours ago
reinventErrors
Node.js 4.3
288 bytes
3 hours ago
Feedback
English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use


Lambda Management Console
CloudWatch Management Con
ALARM: "High 5XX errors"
Secure | https://us-west-2.console.aws.amazon.com/lambda/home?region=us-west-2#/functions/reinventAuthorizer?tab=configuration
Apps
Customer Information
OnCall
K2 Dashboard
Runbook
ES Dashboard
Fleet Management C...
Drive
API Gateway Wiki
Services
Resource Groups
Bob Kinney
Oregon
Support
AWS Lambda
Dashboard
Functions
Lambda > Functions > reinventAuthorizer
ARN - arn:aws:lambda:us-west-2:099174454562:function:reinventAuthorizer
reinventAuthorizer
Qualifiers
Actions
Select a test event..
Test
Configuration
Triggers
Monitoring
Function code
Code entry type
Edit code inline
Runtime
Node.js 4.3
Handler Info
index.handler
index.js
'use strict';
console.log('Loading function');
/**
* AuthPolicy receives a set of allowed and denied methods and generates a valid
* AWS policy for the API Gateway authorizer. The constructor receives the calling
* user principal, the AWS account ID of the API owner, and an apiOptions object.
* The apiOptions can contain an API Gateway RestApi Id, a region for the RestApi, and a
* stage that calls should be allowed/denied for. For example
Feedback
English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

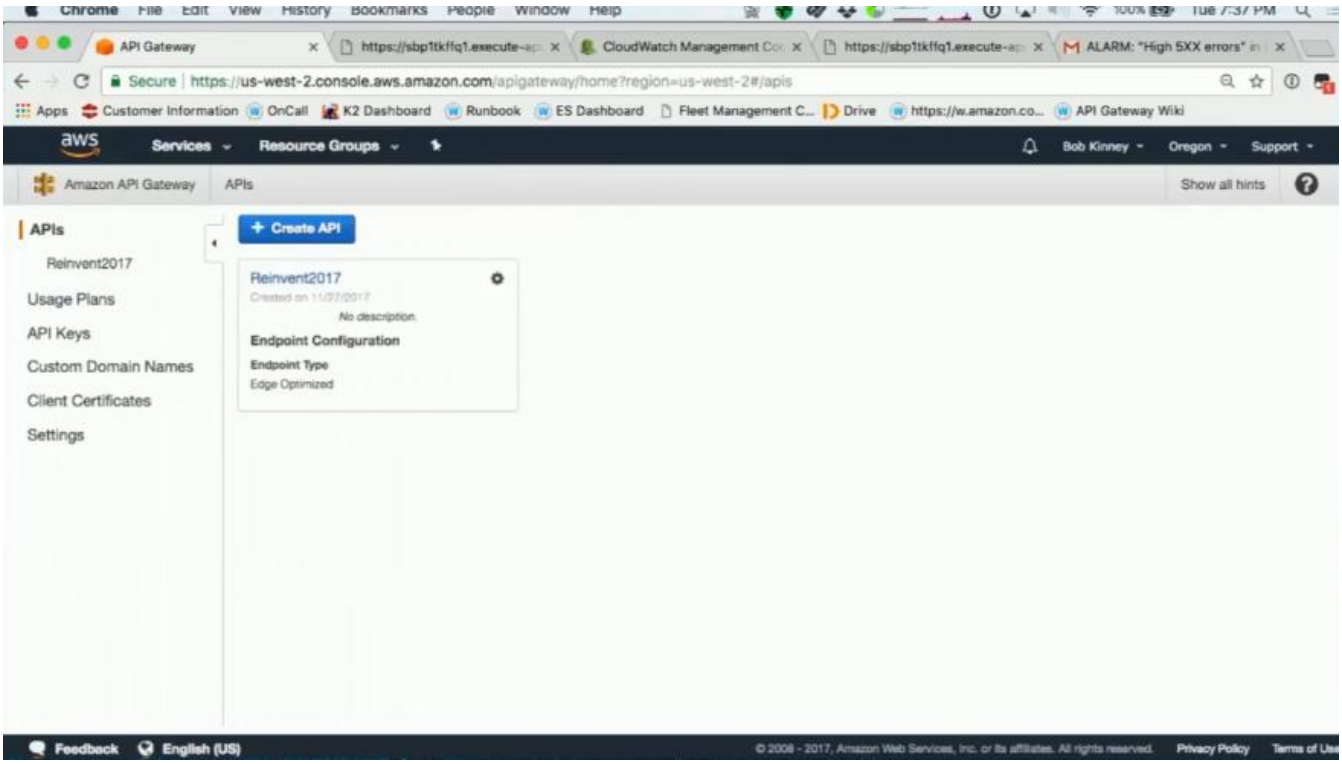




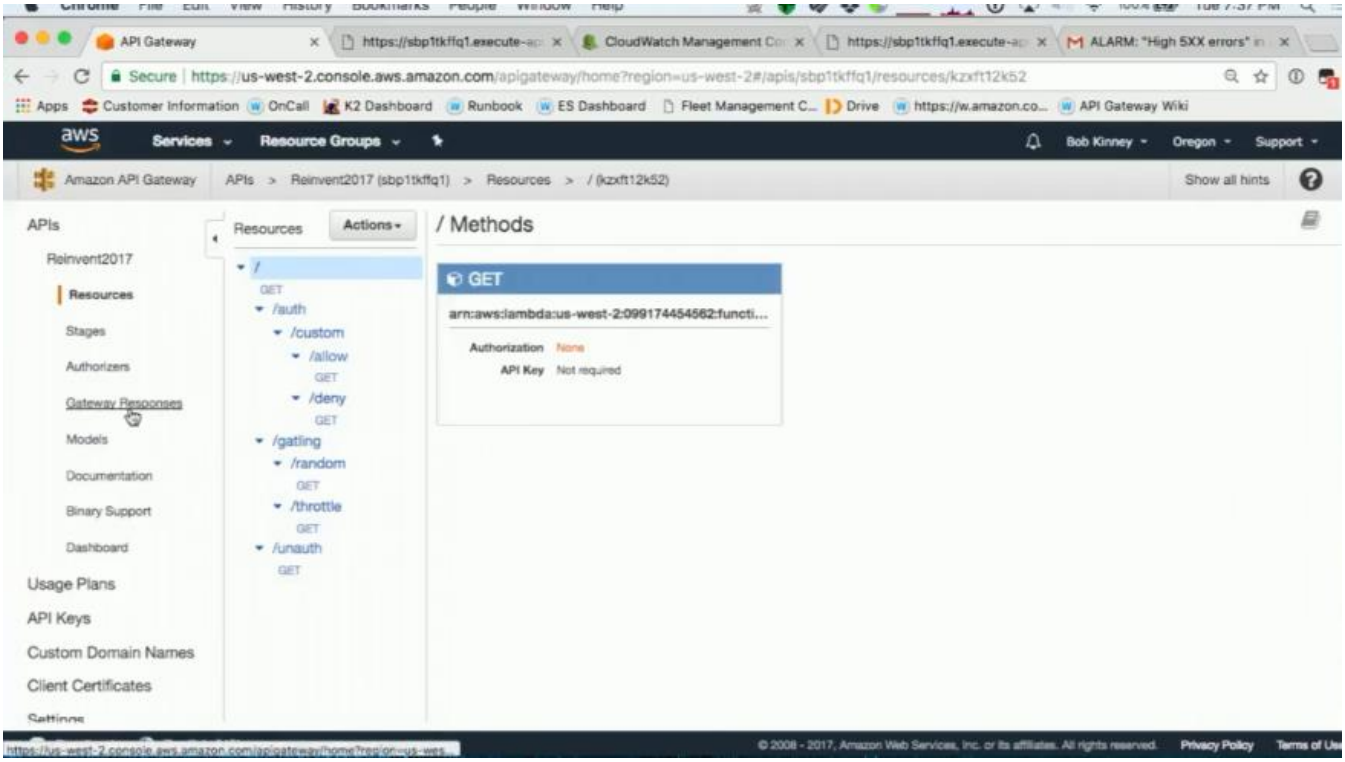
Our request authorizer lambda is going to take the request header, get the authorization header bearer token, and figure out what the username and password is,

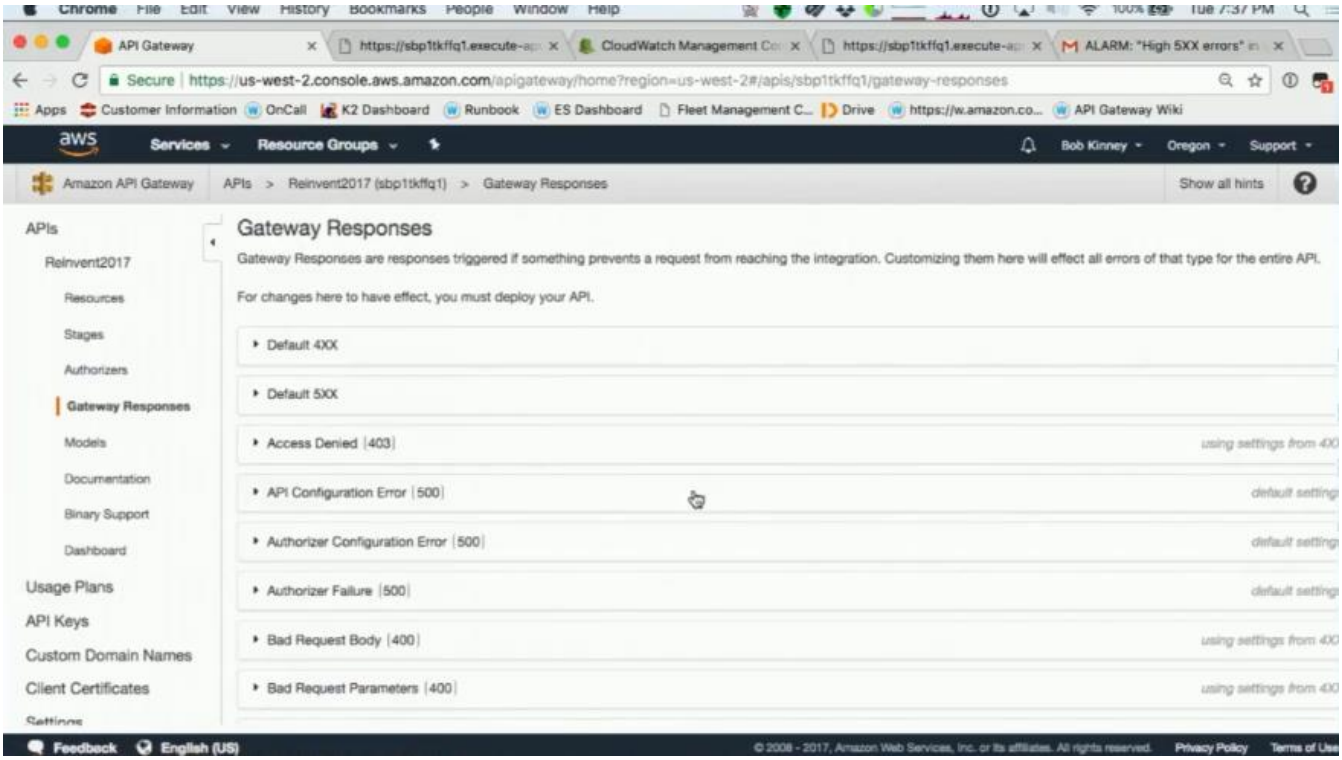


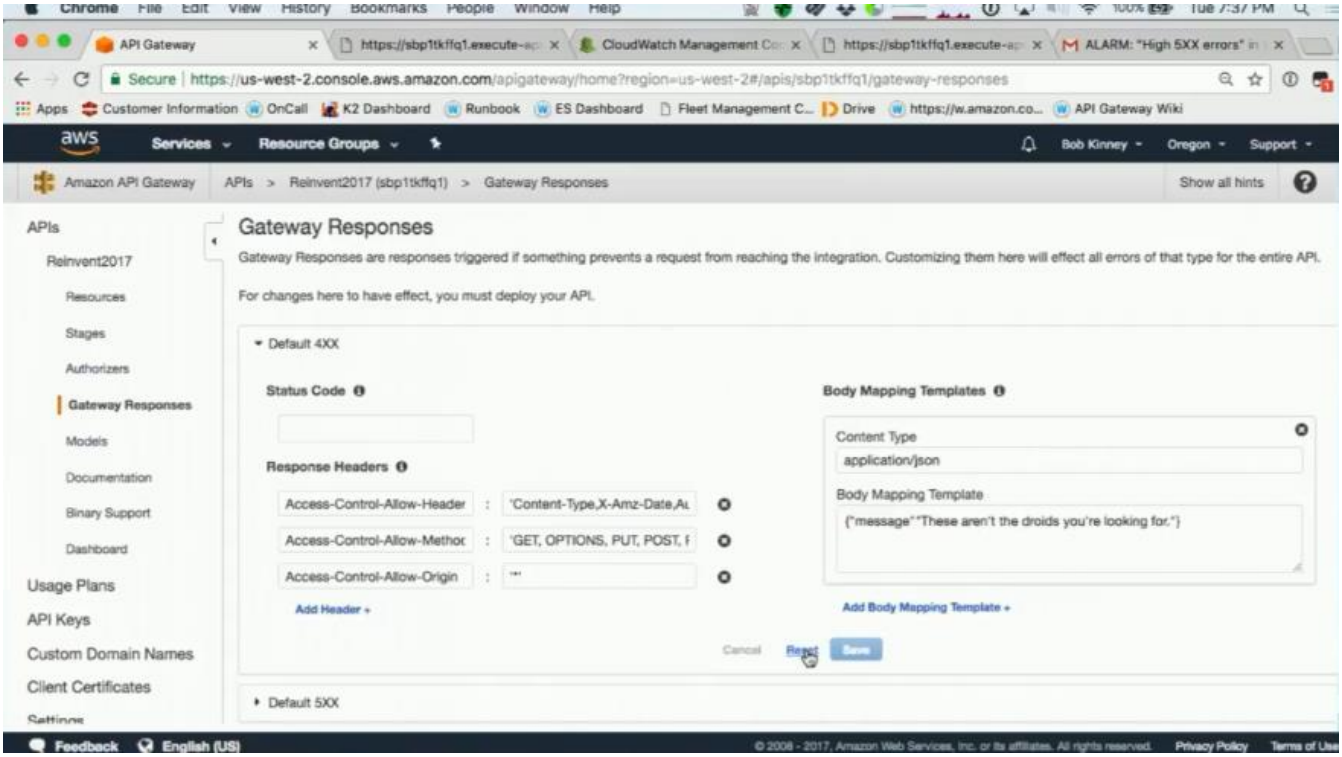
Then compare it against some hard-coded set of credentials as above. We can instead have asked a database for the details like username and password for the token credentials and validate if we want. Above, we will be returning an unauthorized response back for any unauthorized user, this will then lead to the request being rejected.

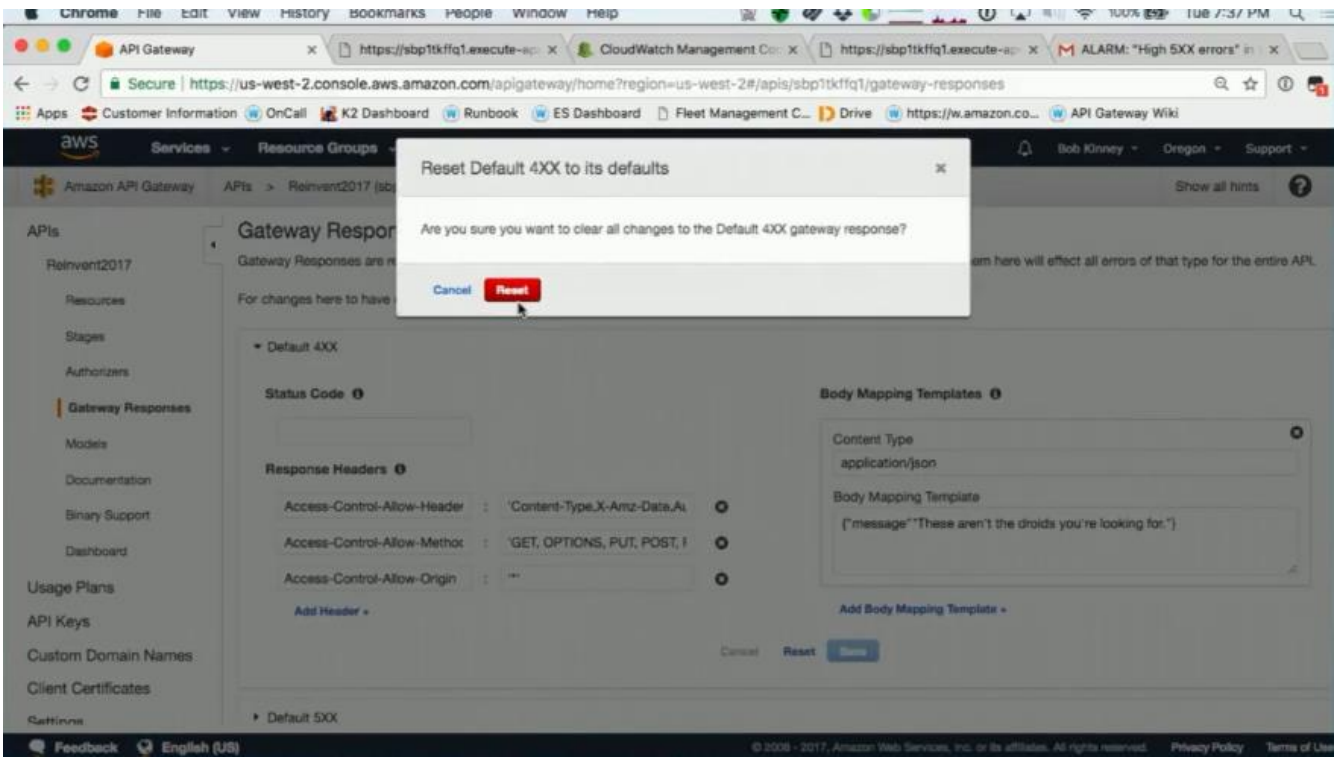


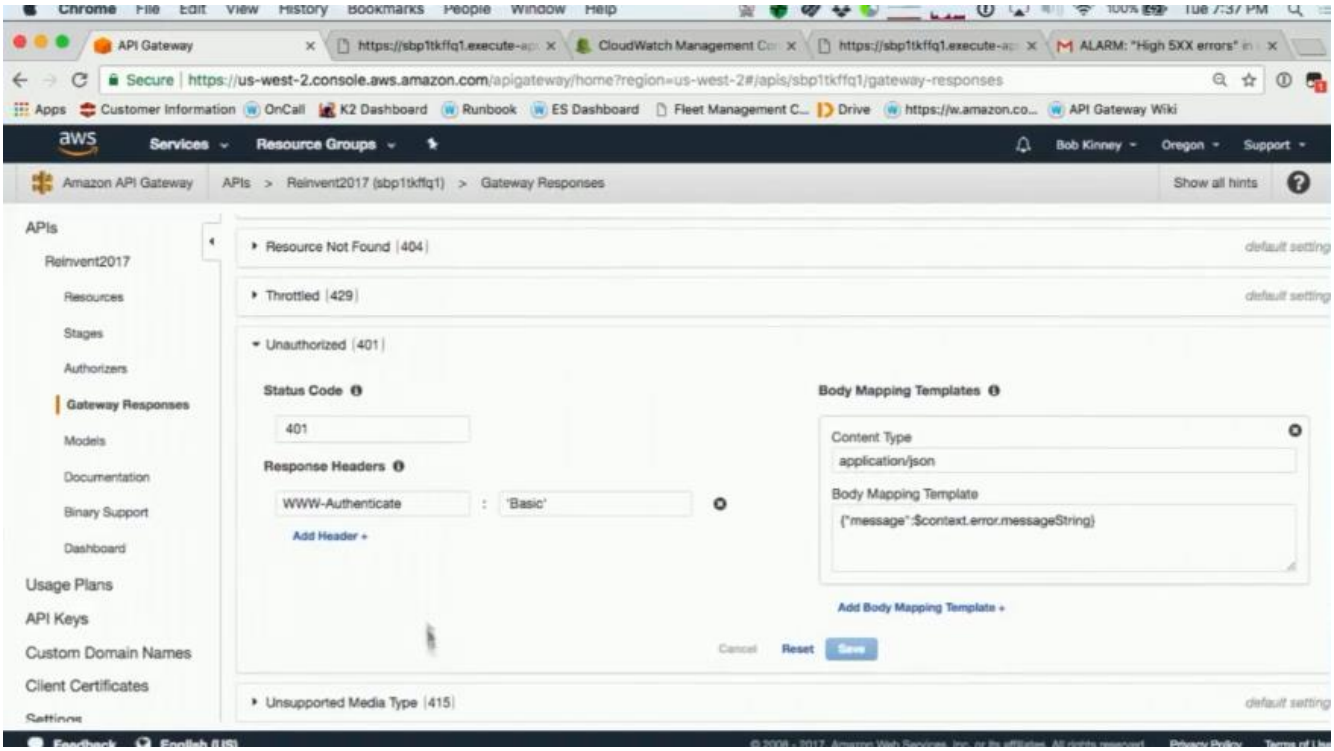
We also need to configure the response we want to give when no authentication is configured or a 401 response,


API Gateway
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/resources/kzxft12k52
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Resources > / (kzxft12k52)
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Resources
Actions
/ Methods
GET
arn:aws:lambda:us-west-2:099174454562:functi...
Authorization None
API Key Not required
/auth
/custom
/allow
/deny
/gatling
/random
/throttle
/unauth
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

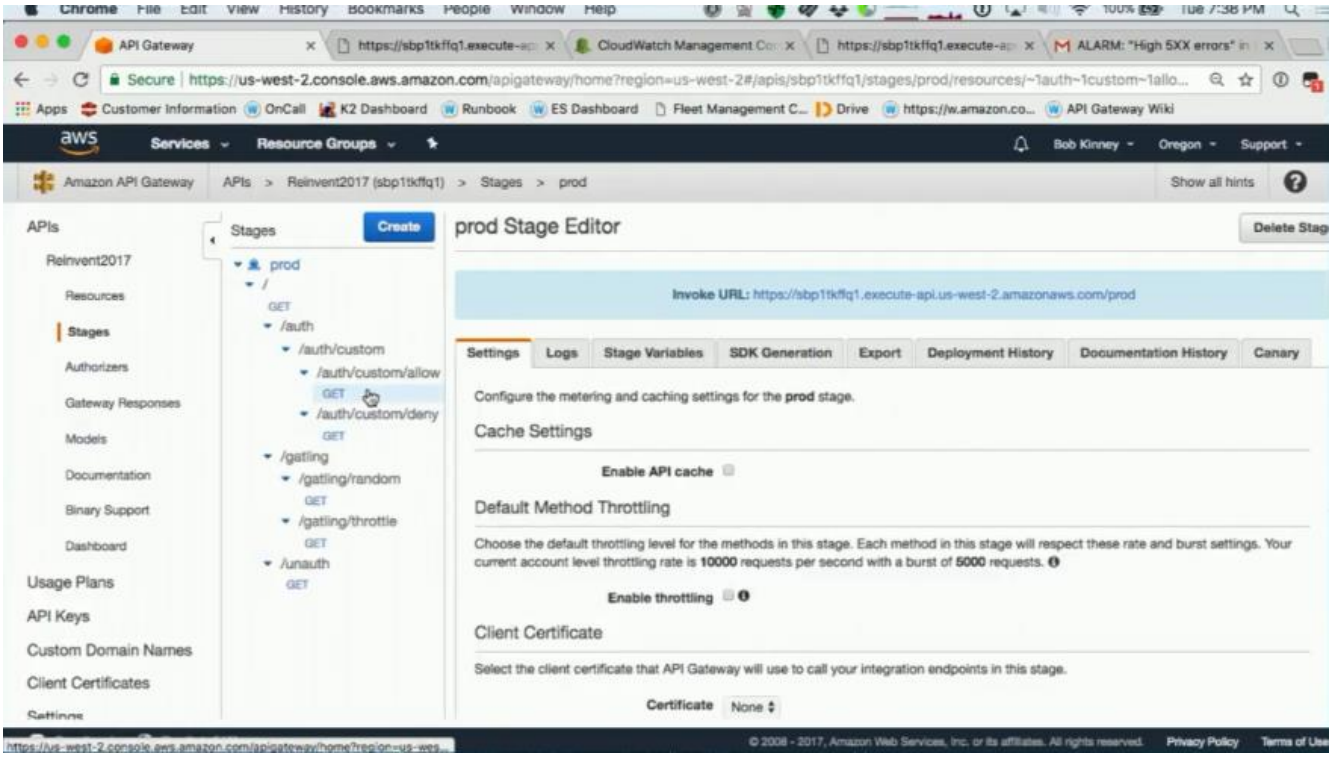

API Gateway
Secure | https://us-west-2.console.aws.amazon.com/apigateway/home?region=us-west-2#/apis/sbp1tkffq1/gateway-responses
Services
Resource Groups
Bob Kinney
Oregon
Support
Amazon API Gateway
APIs > Reinvent2017 (sbp1tkffq1) > Gateway Responses
Show all hints
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates
Gateway Responses
Gateway Responses are responses triggered if something prevents a request from reaching the integration. Customizing them here will effect all errors of that type for the entire API.
For changes here to have effect, you must deploy your API.
Default 4XX
Default 5XX
Access Denied | 403 |
API Configuration Error | 500 |
Authorizer Configuration Error | 500 |
Authorizer Failure | 500 |
Bad Request Body | 400 |
Bad Request Parameters | 400 |
Feedback
English (US)
© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use

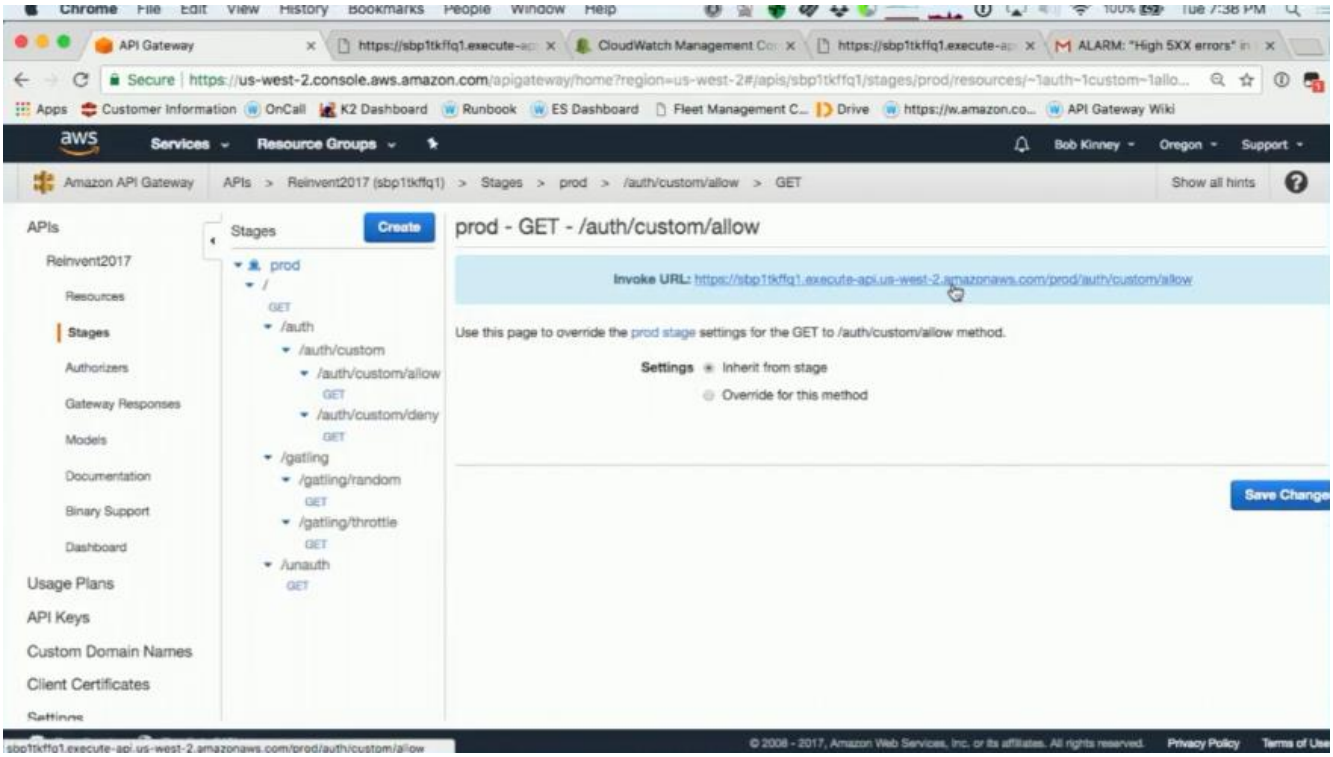

Gateway Responses
Gateway Responses are responses triggered if something prevents a request from reaching the integration. Customizing them here will effect all errors of that type for the entire API.
For changes here to have effect, you must deploy your API.
Default 4XX
Status Code
Response Headers
Access-Control-Allow-Header : 'Content-Type,X-Amz-Date,A
Access-Control-Allow-Method : 'GET, OPTIONS, PUT, POST,
Access-Control-Allow-Origin : '*'
Add Header +
Body Mapping Templates
Content Type
application/json
Body Mapping Template
{"message":"These aren't the droids you're looking for."}
Add Body Mapping Template +
Cancel
Reset
Save
Default 5XX
APIs
Reinvent2017
Resources
Stages
Authorizers
Gateway Responses
Models
Documentation
Binary Support
Dashboard
Usage Plans
API Keys
Custom Domain Names
Client Certificates

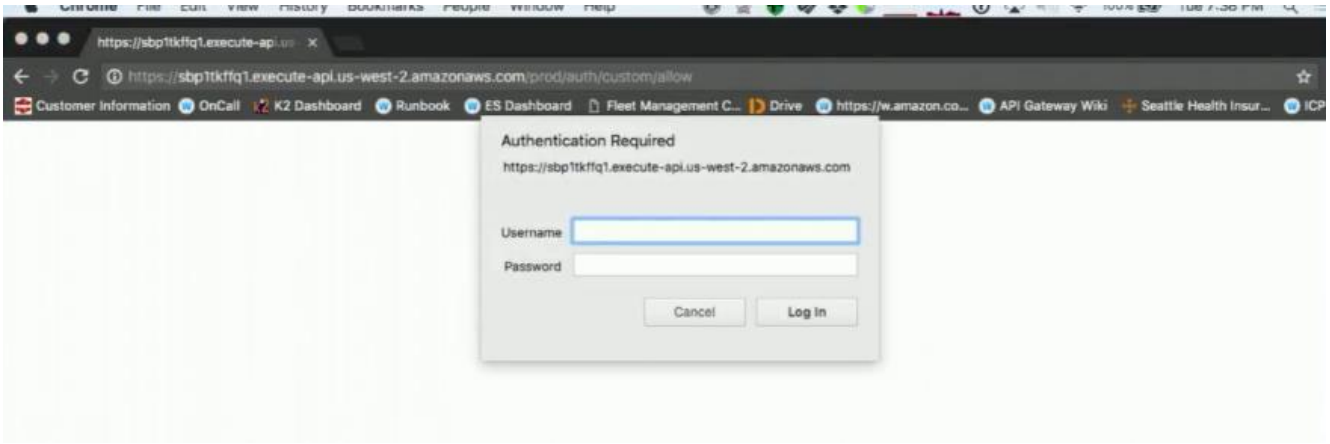

Reset Default 4XX to its defaults
Are you sure you want to clear all changes to the Default 4XX gateway response?
Cancel
Reset



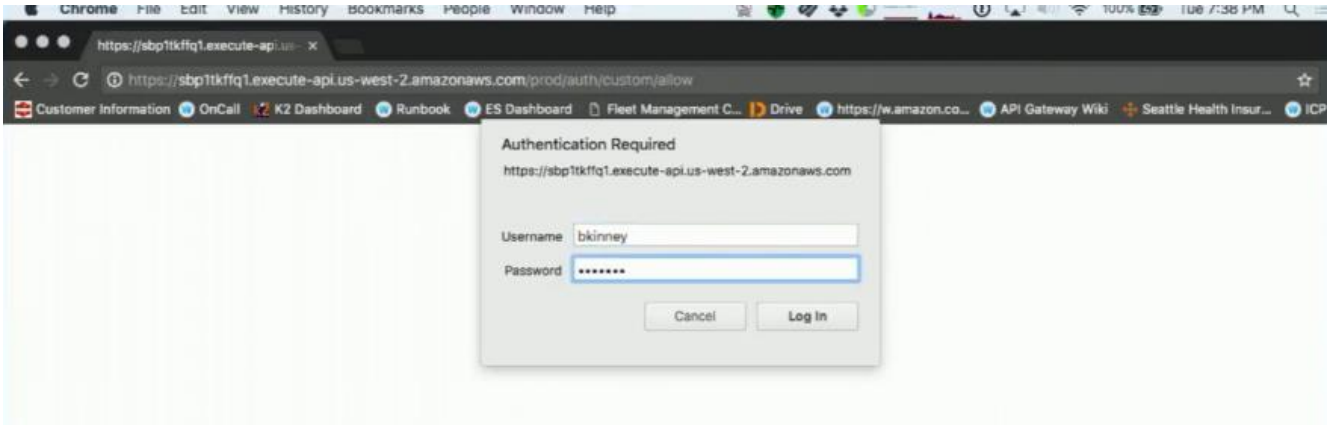
We add the 'WWW-Authenticate' header with the value 'Basic'. This means that when we hit this API without credentials configured, we should get a 401 response back

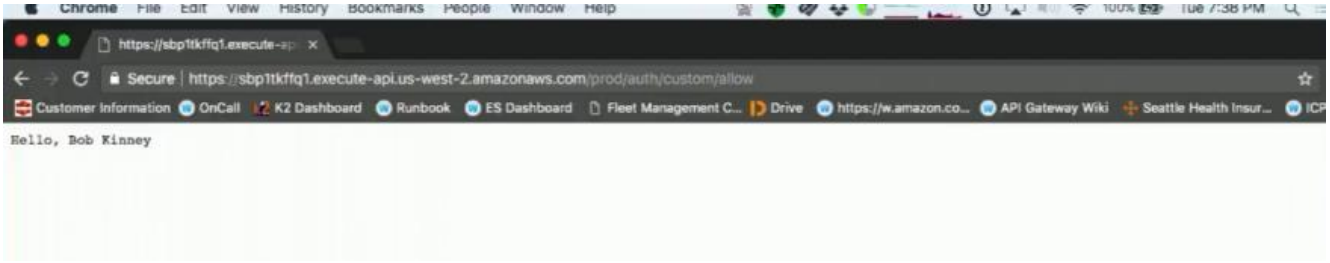




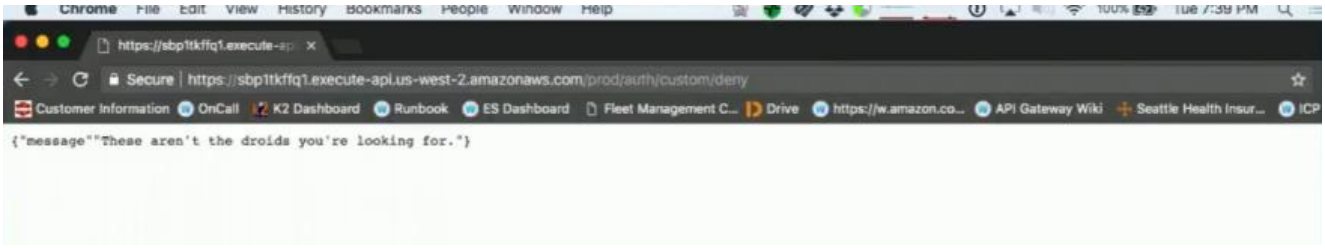


As you can see when we hit this API having basic auth, we get the 401 response back and the browser prompts us to enter the username and password details.

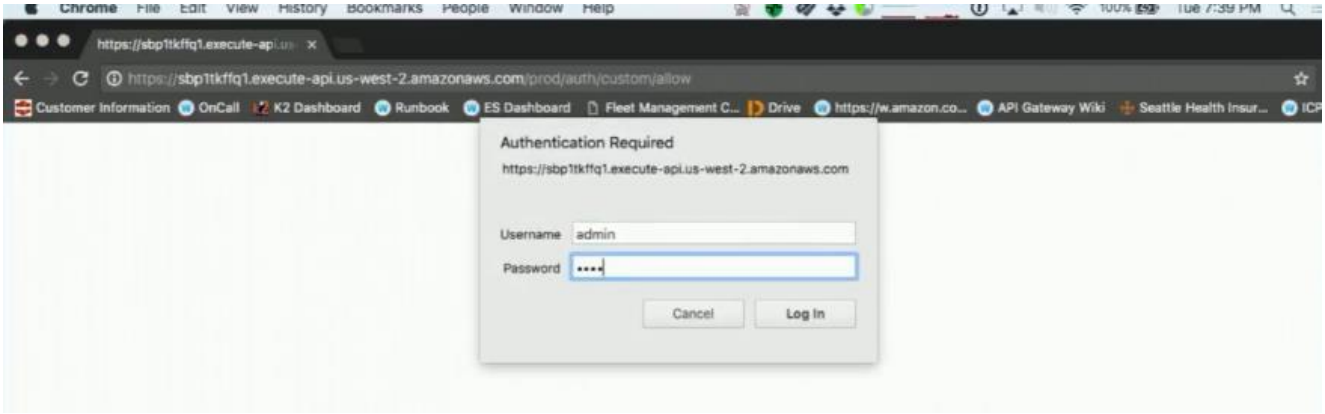




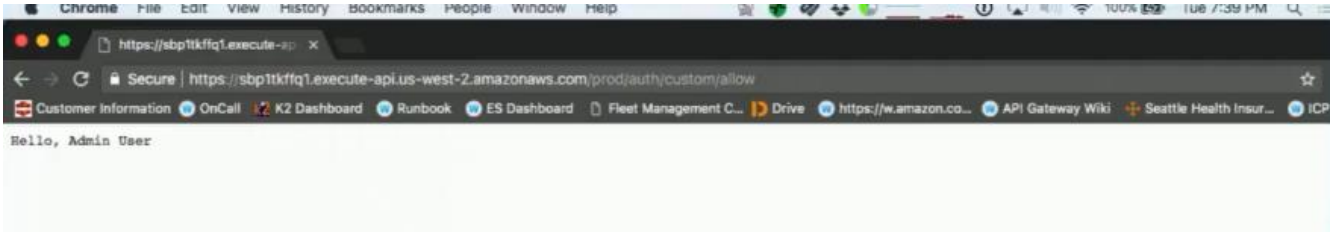
We then can get the correct response having providing the details. This can use any contextual key/value information that we can pass back in the custom authorizer. Note that this user is only allowed access to this specific method, if we try to access other privileged APIs like below, we should get an unauthorized response back



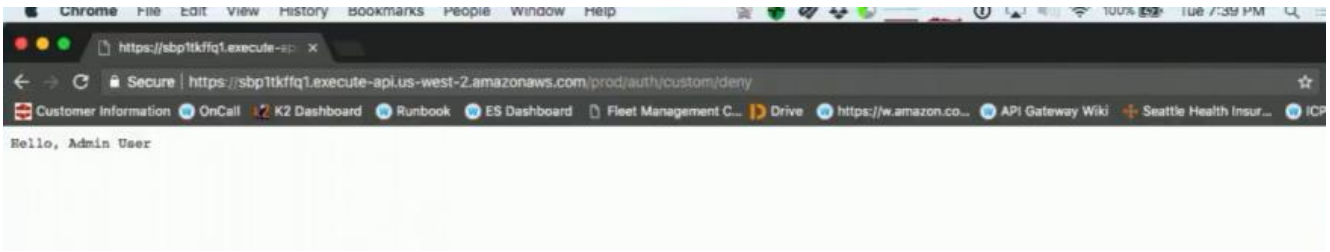
We get the default error message in this case showing that we don't have access to this method



We can again log in as a privileged user as above



We then get logged in and able to see all API endpoints. This is because the authorizer has sent back the contextual details that we are logged in as the admin user



We also can see the privileged resource also. We are now able to protect portions of our APIs from non-privileged users

Check in with Doug

TAMPR promotions have been a hit, and the app is more popular than ever. Doug is now speaking with other services, such as a new site focused on brunch spots, on how they can work together.

Doug wants a way he can expose portions of his API to these third parties, but track their usage for potential billing opportunities.

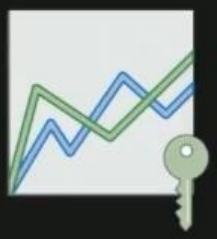
Protecting Your API: Usage Plans

API Gateway usage plans

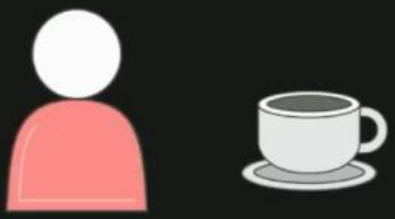
API Key Throttling
- Rate/Burst per API Key

API Key Quota
- Periodic limits per API Key

API Key Usage
- Daily usage records



Check in with Doug



TAMPR is continuing to grow and some design decisions Doug made in the early versions are causing scaling issues.

Doug has rewritten and optimized a large portion of the backend, but he's worried that despite automated tests passing that there may be issues switching 100% of the TAMPR traffic to the new code.

He needs a way to test his new code with a portion of his live traffic.

Managing Your API: Canary Release Deployments

Canary release deployments ★ NEW ★

Tag a new deployment as a "canary" for a stage

- Control the percentage of traffic the canary receives

Isolated metrics and logging

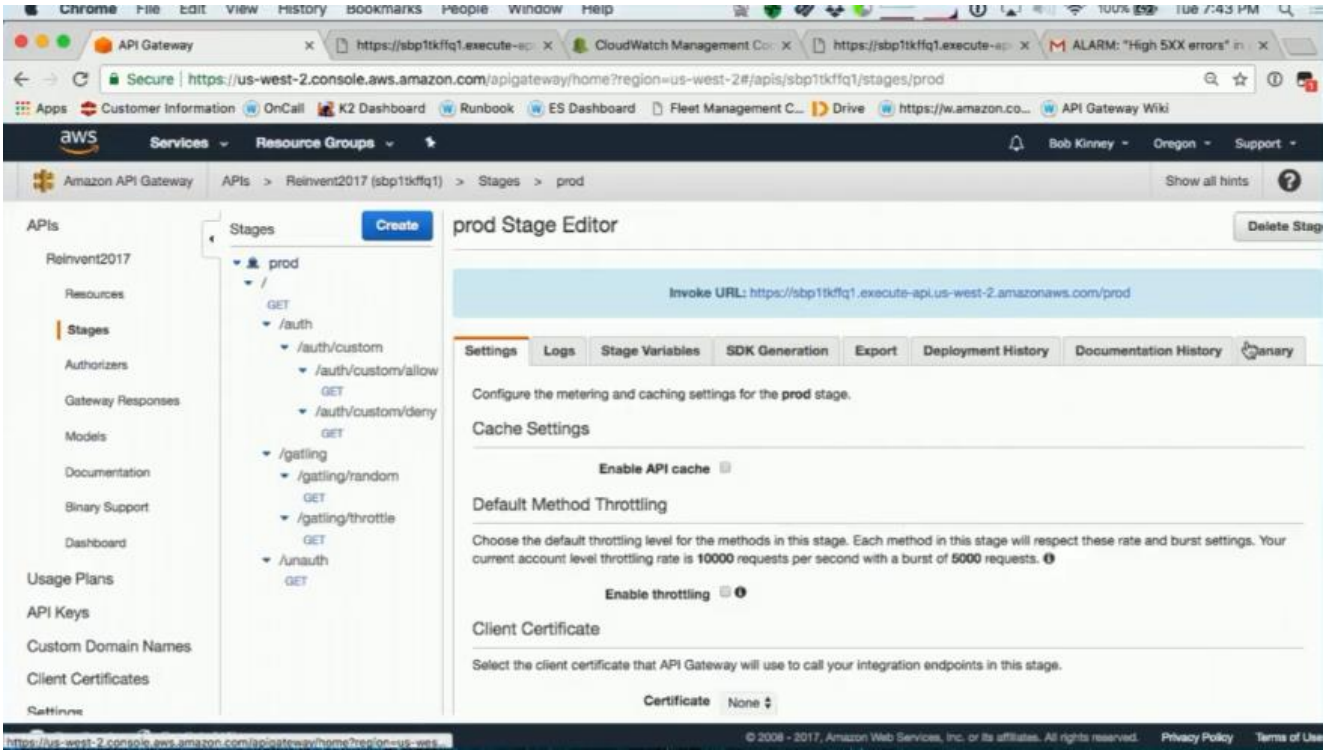
- STAGE/canary

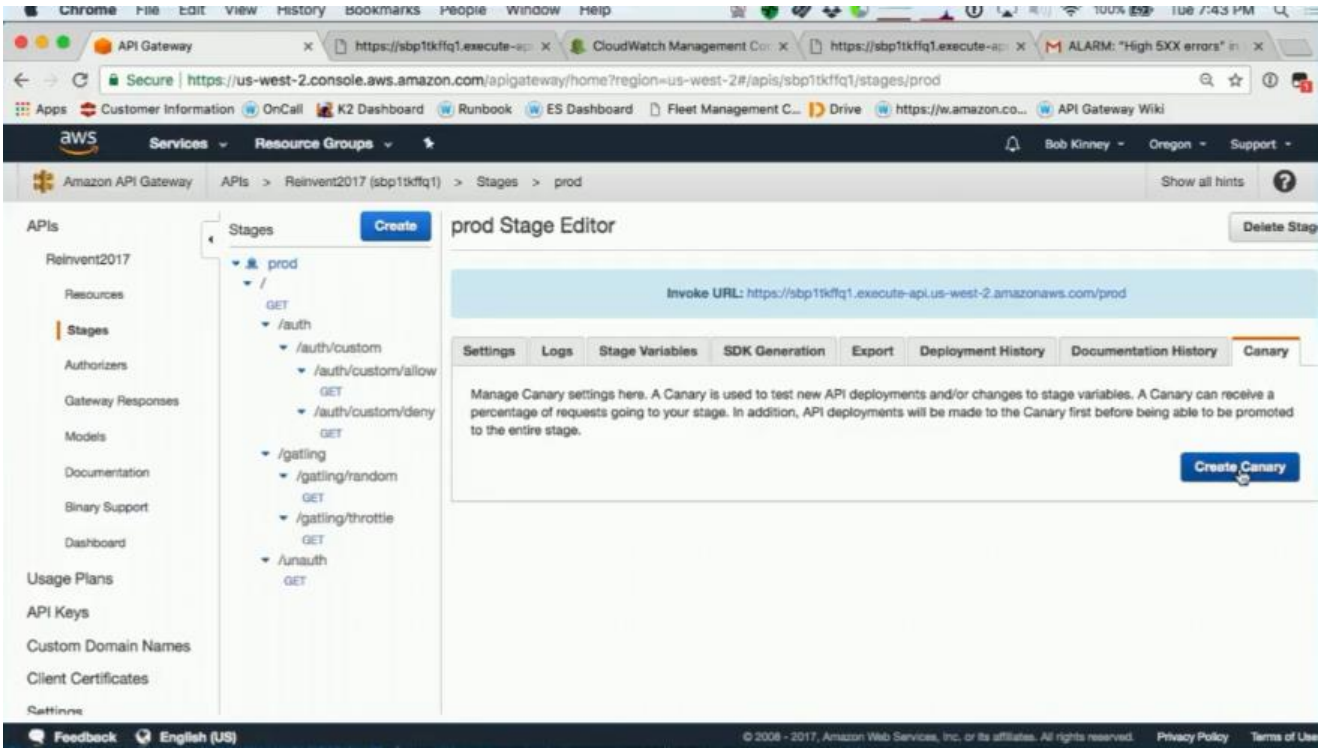
Any changes that can be tied to a deployment can be tested

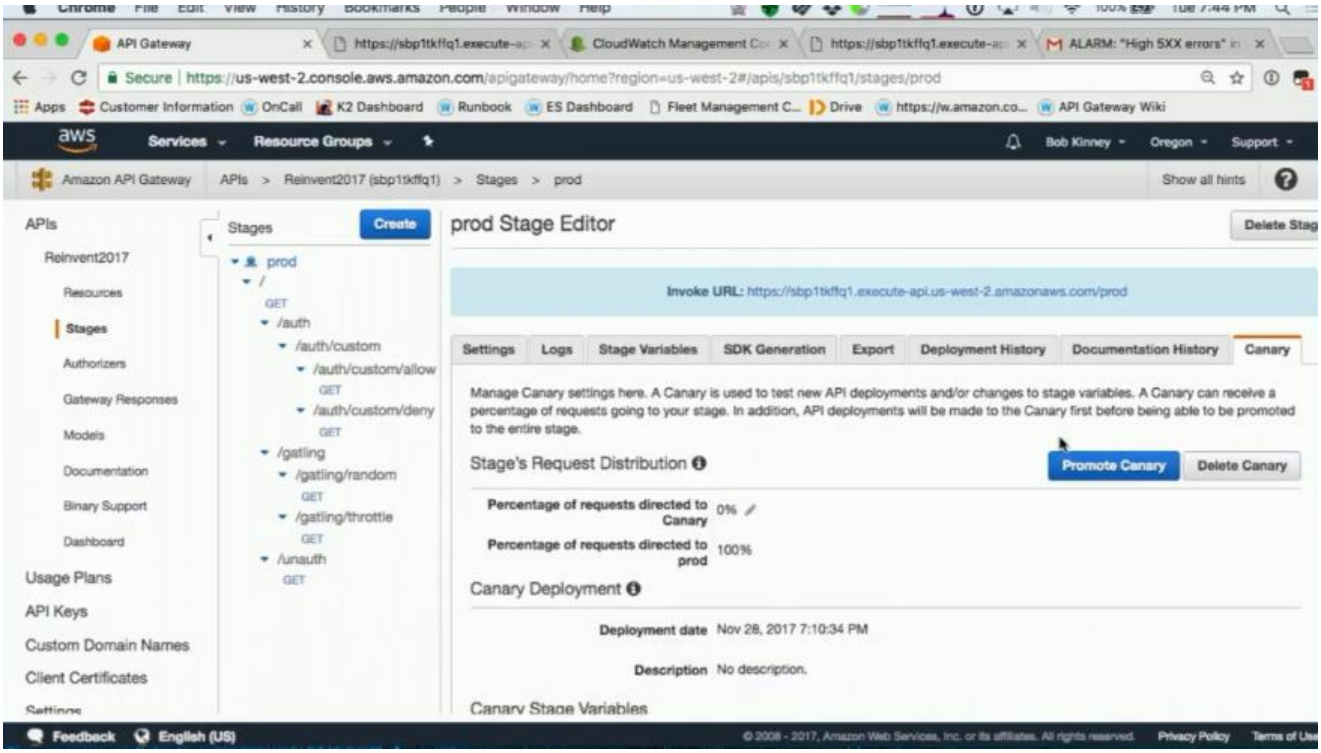
- Including stage variables

Demo

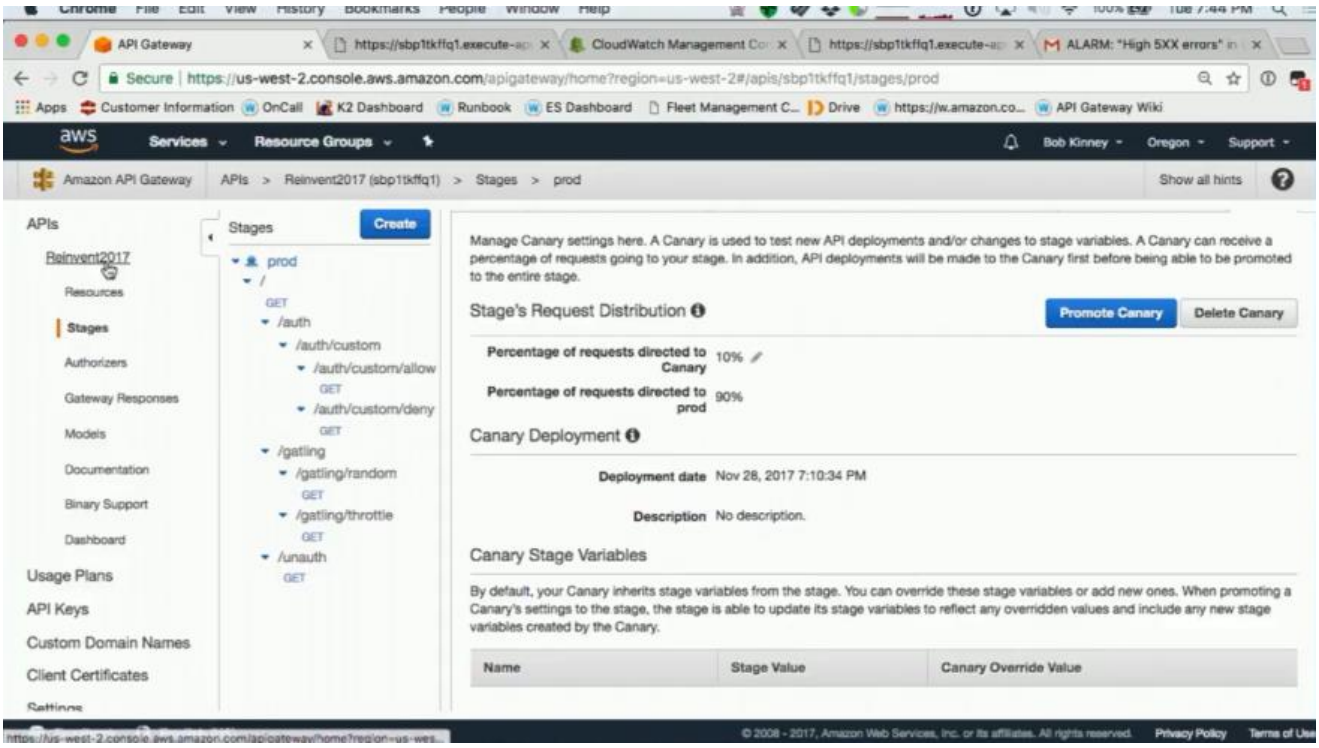
Let us see a demo of using canary deployments



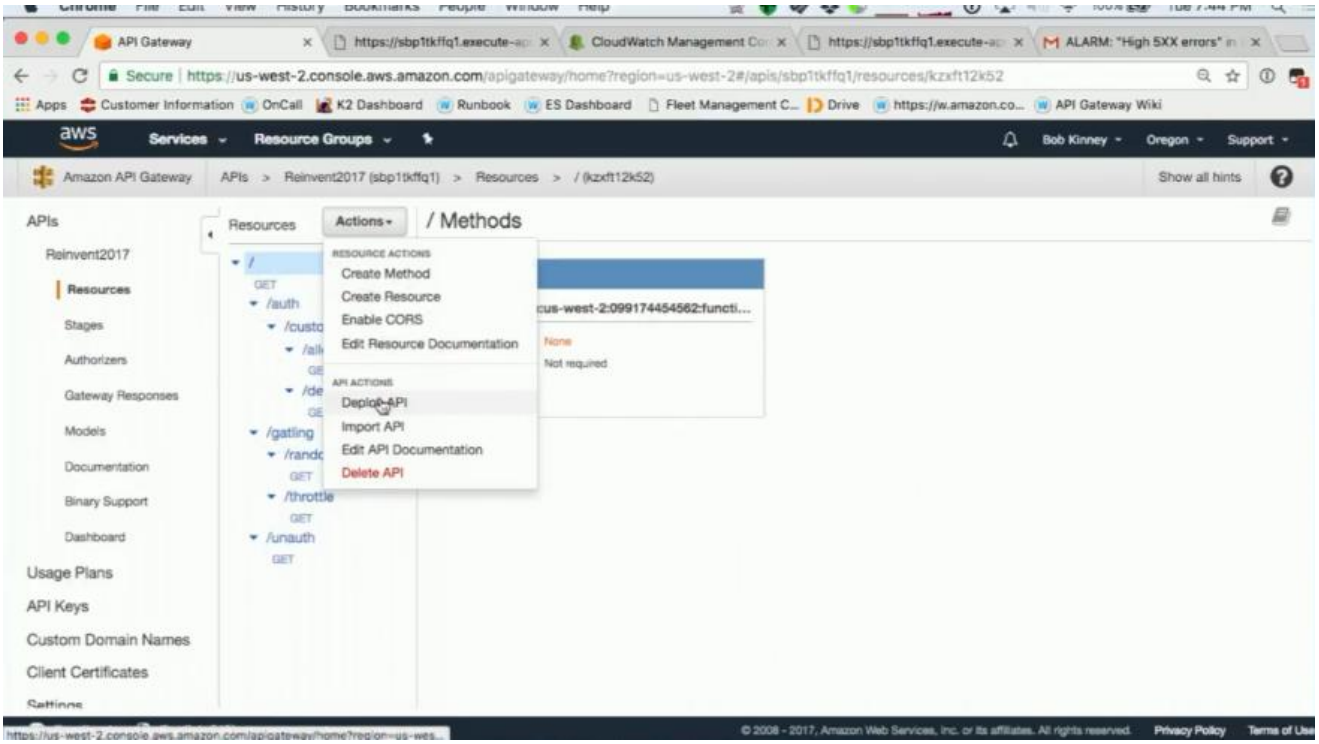


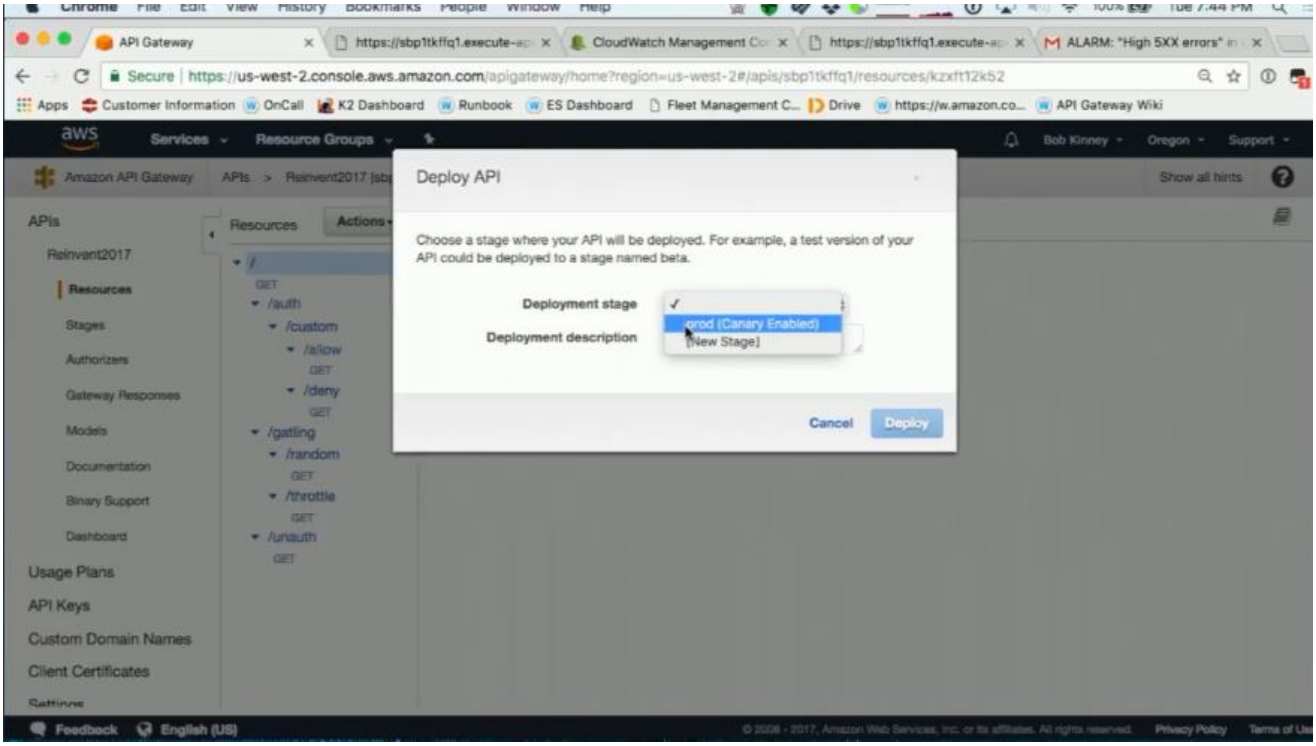


When we create a canary, API Gateway sets up the same deployment that is set up as ‘prod’ in this case as the canary.

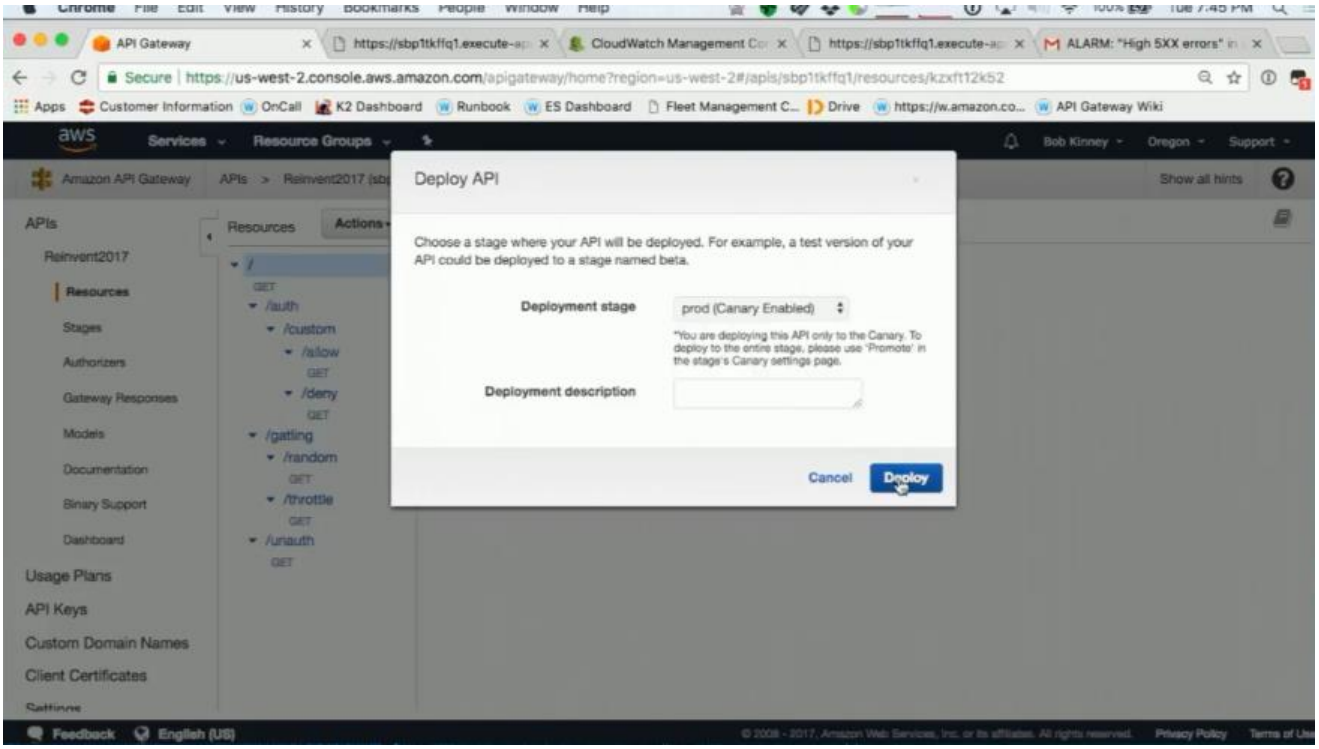


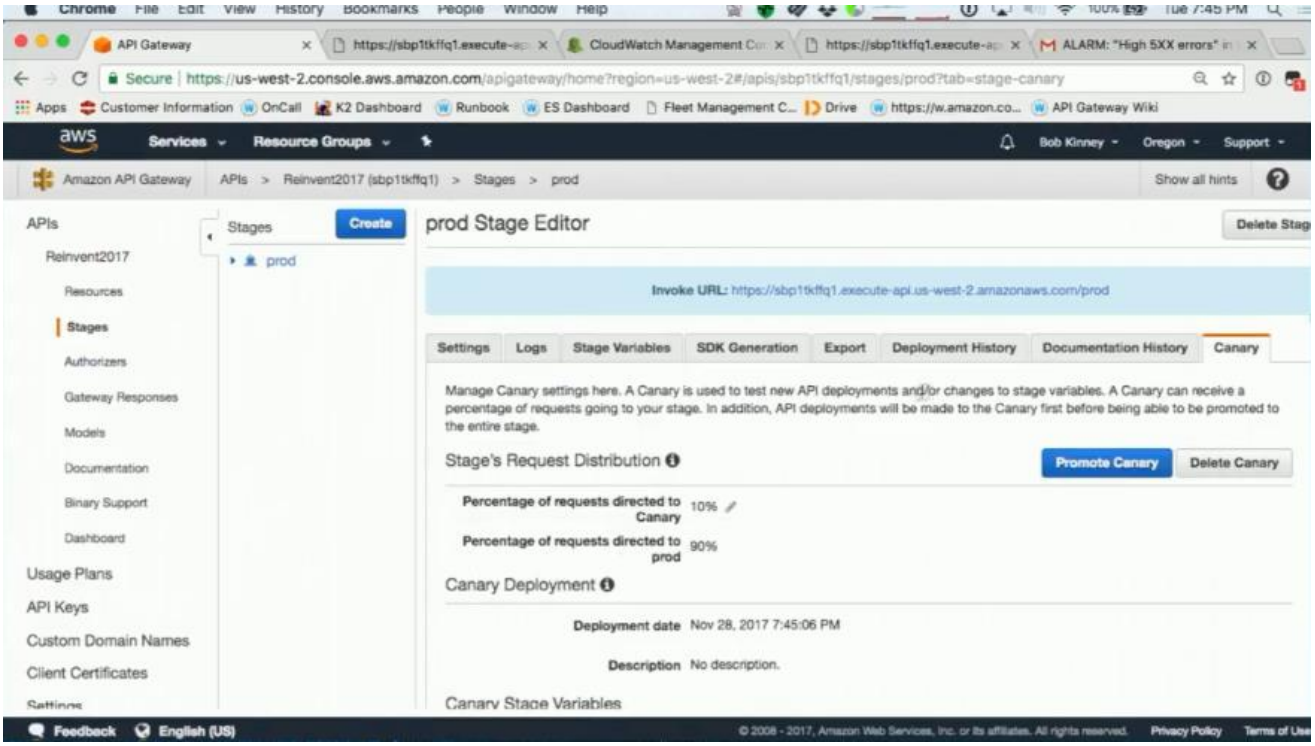
We then set the amount of traffic that we want to sent to our canary like 10% above



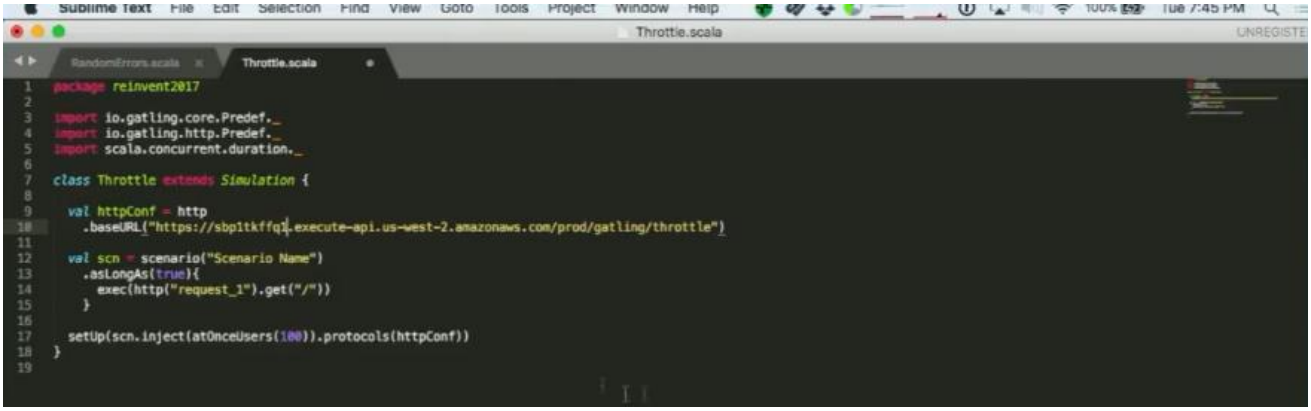


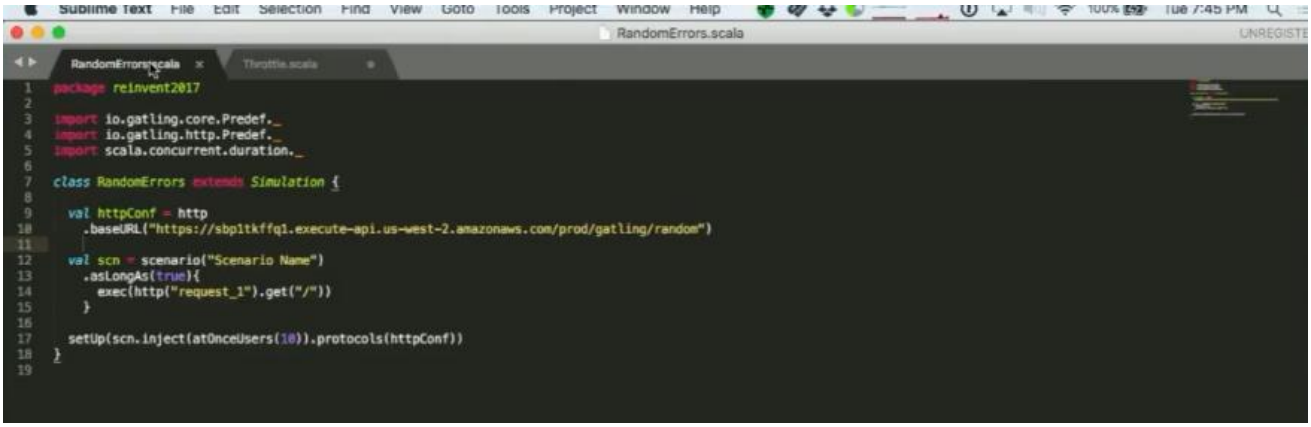
The console by default will force you to deploy to your canary first, then you can promote your canary up to the prod stage after you have validated it is good. The API Gateway also can allow you to bypass this initial stage and just promote your canary if you want, but this 2-stage approach is the recommended way.

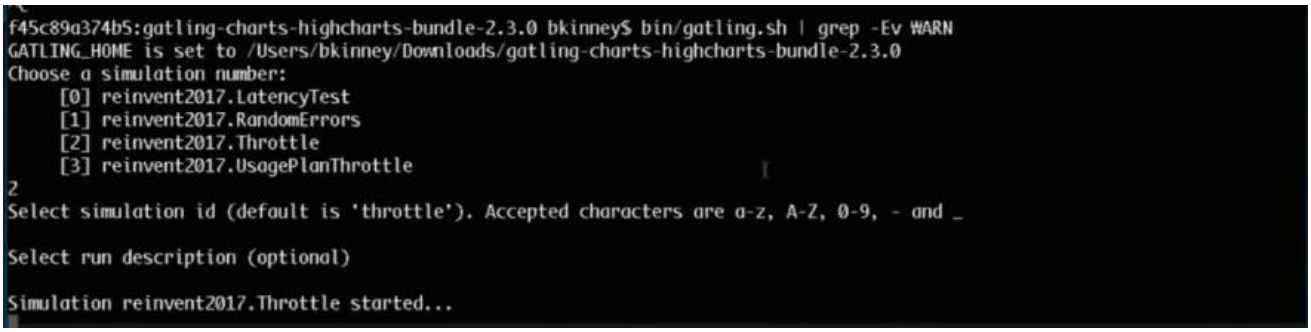


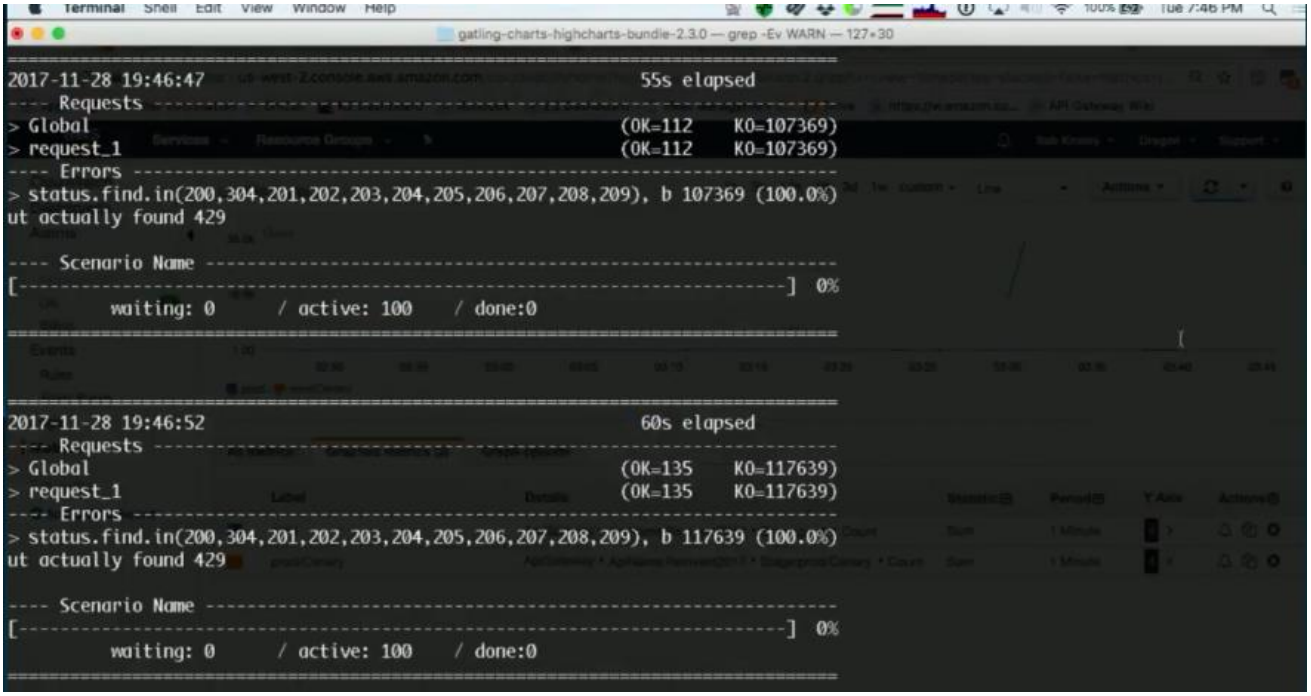


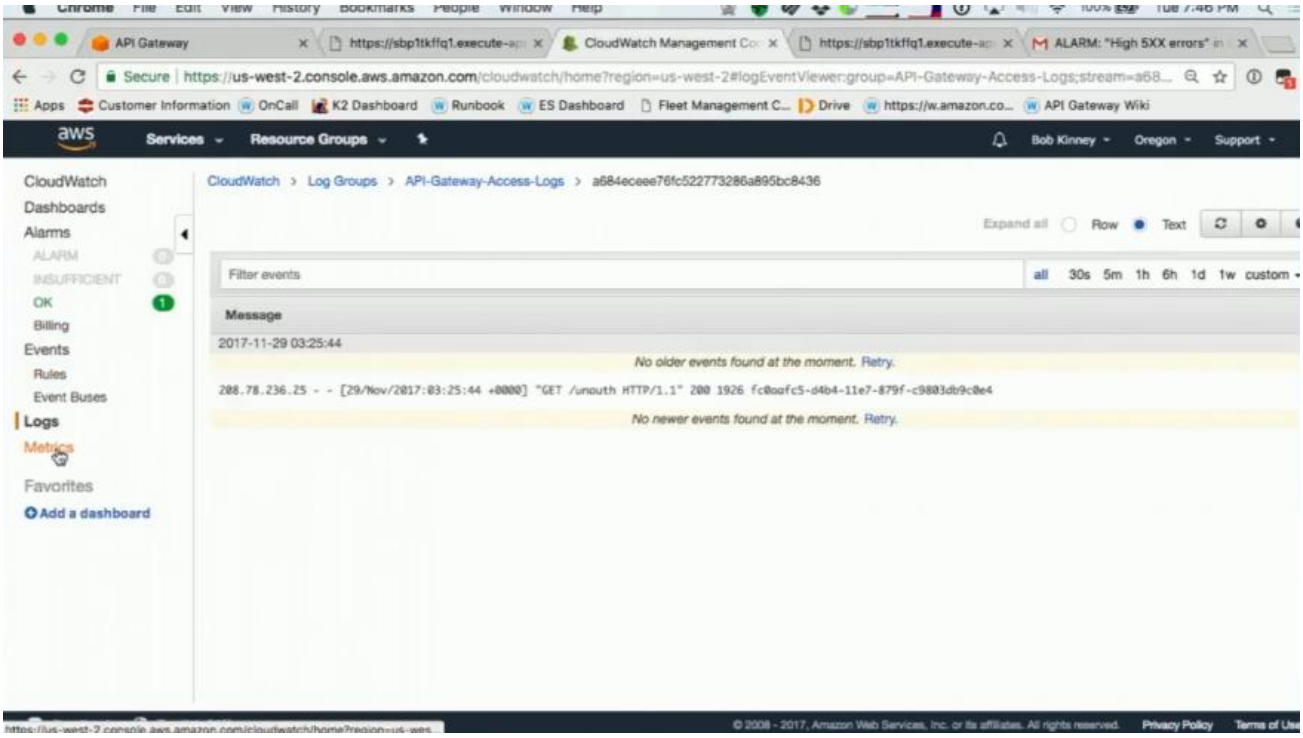
Once we deploy our canary, we can then start up our Gatlin calling code to test as below

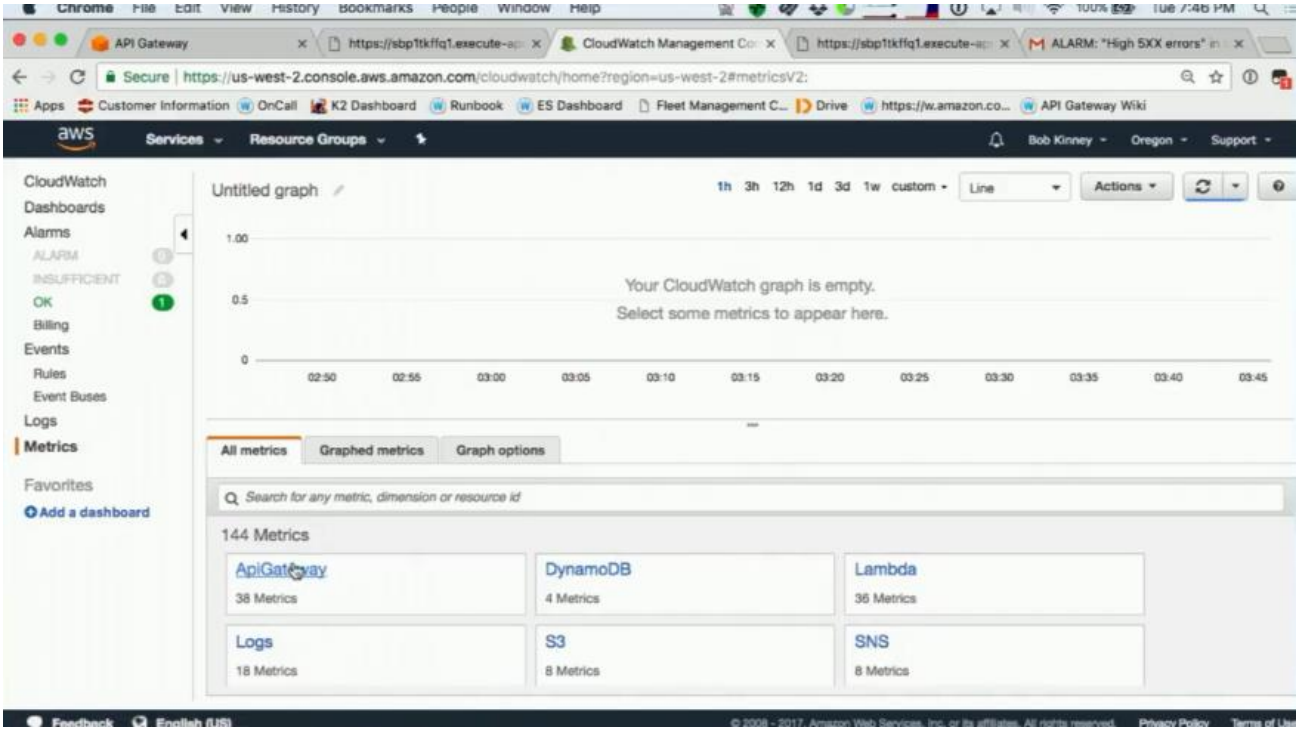

Throttle.scala
RandomErrors.scala
Throttle.scala
package reinvent2017
import io.gatling.core.Predef._
import io.gatling.http.Predef._
import scala.concurrent.duration._
class Throttle extends Simulation {
val httpConf = http
.baseURL("https://sbp1tkffq1.execute-api.us-west-2.amazonaws.com/prod/gatling/throttle")
val scn = scenario("Scenario Name")
.asLongAs(true){
exec(http("request_1").get("/"))
}
setUp(scn.inject(atOnceUsers(100)).protocols(httpConf))
}

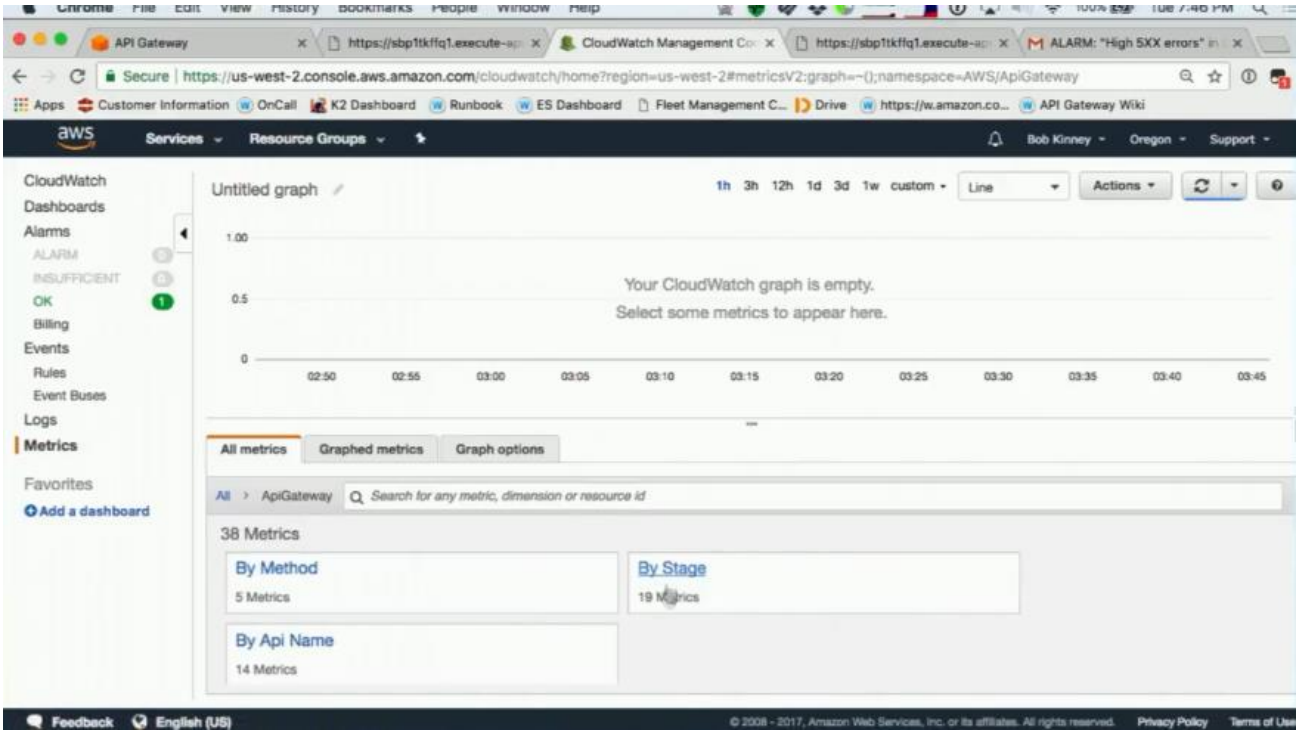

RandomErrors.scala
RandomErrors.scala
Throttle.scala
package reinvent2017
import io.gatling.core.Predef._
import io.gatling.http.Predef._
import scala.concurrent.duration._
class RandomErrors extends Simulation {
val httpConf = http
.baseURL("https://sbp1tkffq1.execute-api.us-west-2.amazonaws.com/prod/gatling/random")
val scn = scenario("Scenario Name")
.asLongAs(true){
exec(http("request_1").get("/"))
}
setUp(scn.inject(atOnceUsers(10)).protocols(httpConf))
}

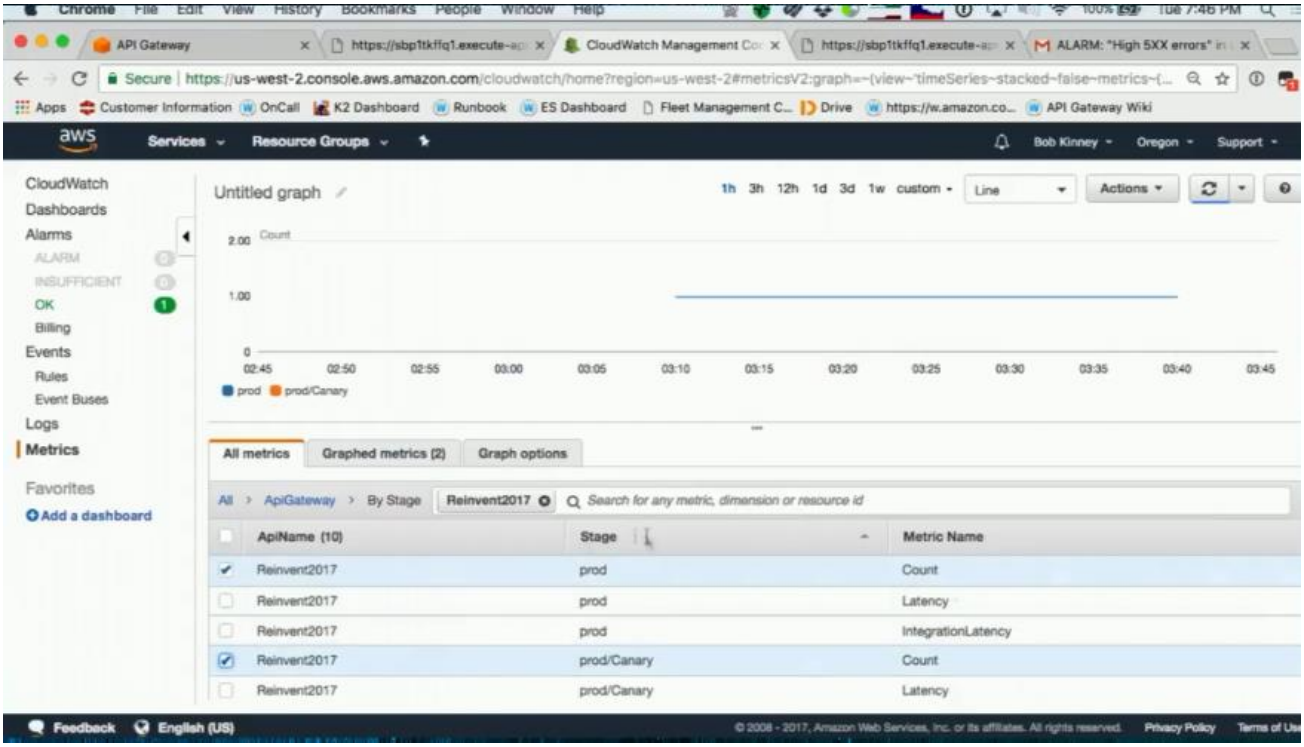

f45c89a374b5:gatling-charts-highcharts-bundle-2.3.0 bkinney$ bin/gatling.sh | grep -Ev WARN
GATLING_HOME is set to /Users/bkinney/Downloads/gatling-charts-highcharts-bundle-2.3.0
Choose a simulation number:
[0] reinvent2017.LatencyTest
[1] reinvent2017.RandomErrors
[2] reinvent2017.Throttle
[3] reinvent2017.UsagePlanThrottle
2
Select simulation id (default is 'throttle'). Accepted characters are a-z, A-Z, 0-9, - and _
Select run description (optional)
Simulation reinvent2017.Throttle started...


gatling-charts-highcharts-bundle-2.3.0 — grep -Ev WARN — 127×30
2017-11-28 19:46:47 55s elapsed
---- Requests
> Global (OK=112 KO=107369)
> request_1 (OK=112 KO=107369)
---- Errors
> status.find.in(200,304,201,202,203,204,205,206,207,208,209), b 107369 (100.0%)
ut actually found 429
---- Scenario Name
[] 0%
waiting: 0 / active: 100 / done:0
2017-11-28 19:46:52 60s elapsed
---- Requests
> Global (OK=135 KO=117639)
> request_1 (OK=135 KO=117639)
---- Errors
> status.find.in(200,304,201,202,203,204,205,206,207,208,209), b 117639 (100.0%)
ut actually found 429
---- Scenario Name
[] 0%
waiting: 0 / active: 100 / done:0

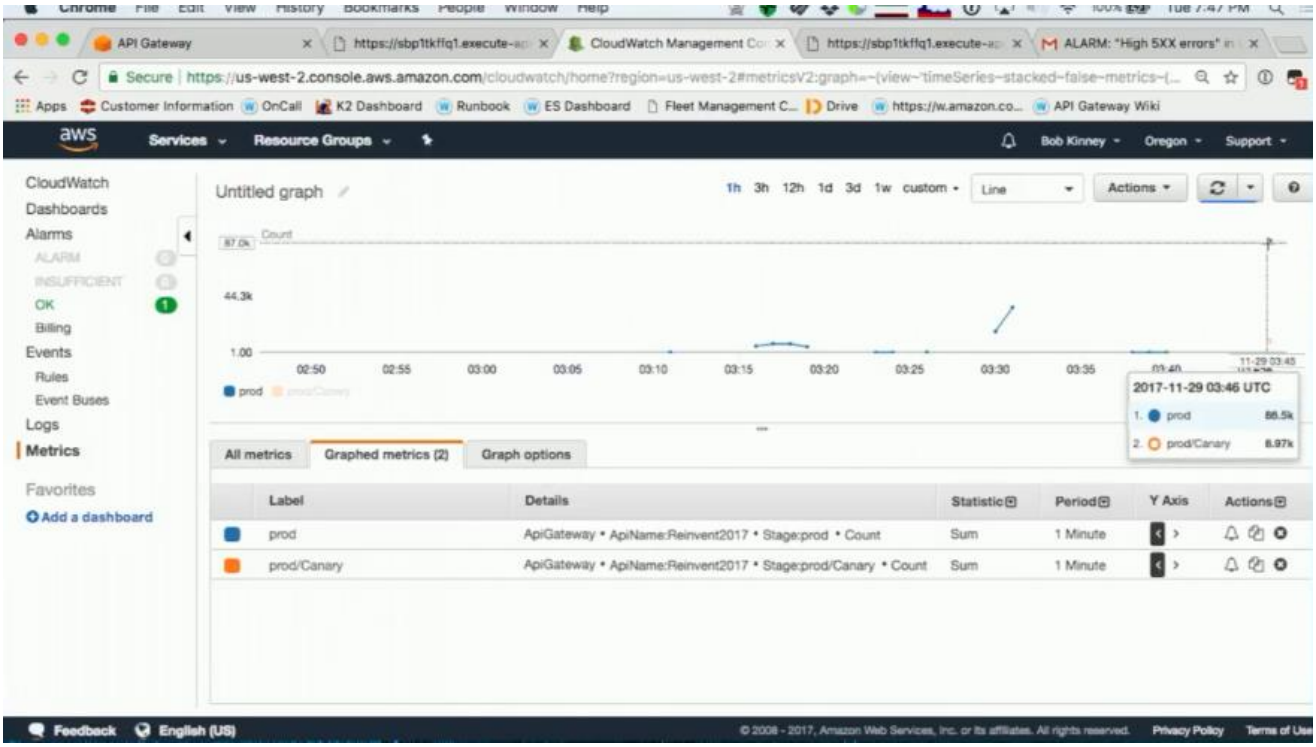

CloudWatch
Dashboards
Alarms
ALARM
INSUFFICIENT
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
CloudWatch > Log Groups > API-Gateway-Access-Logs > a684eceee76fc522773286a895bc8436
Filter events
Message
2017-11-29 03:25:44
No older events found at the moment. Retry.
No newer events found at the moment. Retry.


CloudWatch
Dashboards
Alarms
OK
Billing
Events
Rules
Event Buses
Logs
Metrics
Favorites
Add a dashboard
Untitled graph
Your CloudWatch graph is empty.
Select some metrics to appear here.
All metrics
Graphed metrics
Graph options
Search for any metric, dimension or resource id
144 Metrics
ApiGateway
38 Metrics
DynamoDB
4 Metrics
Lambda
36 Metrics
Logs
18 Metrics
S3
8 Metrics
SNS
8 Metrics

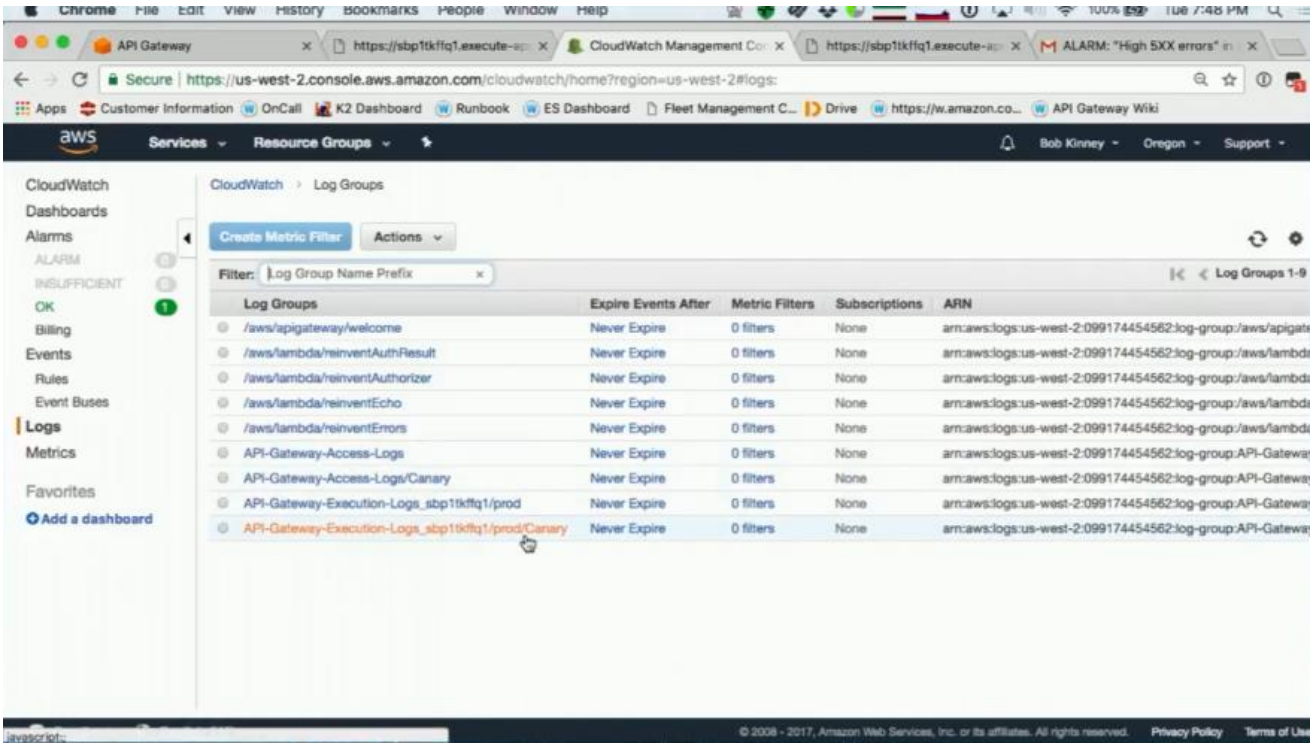


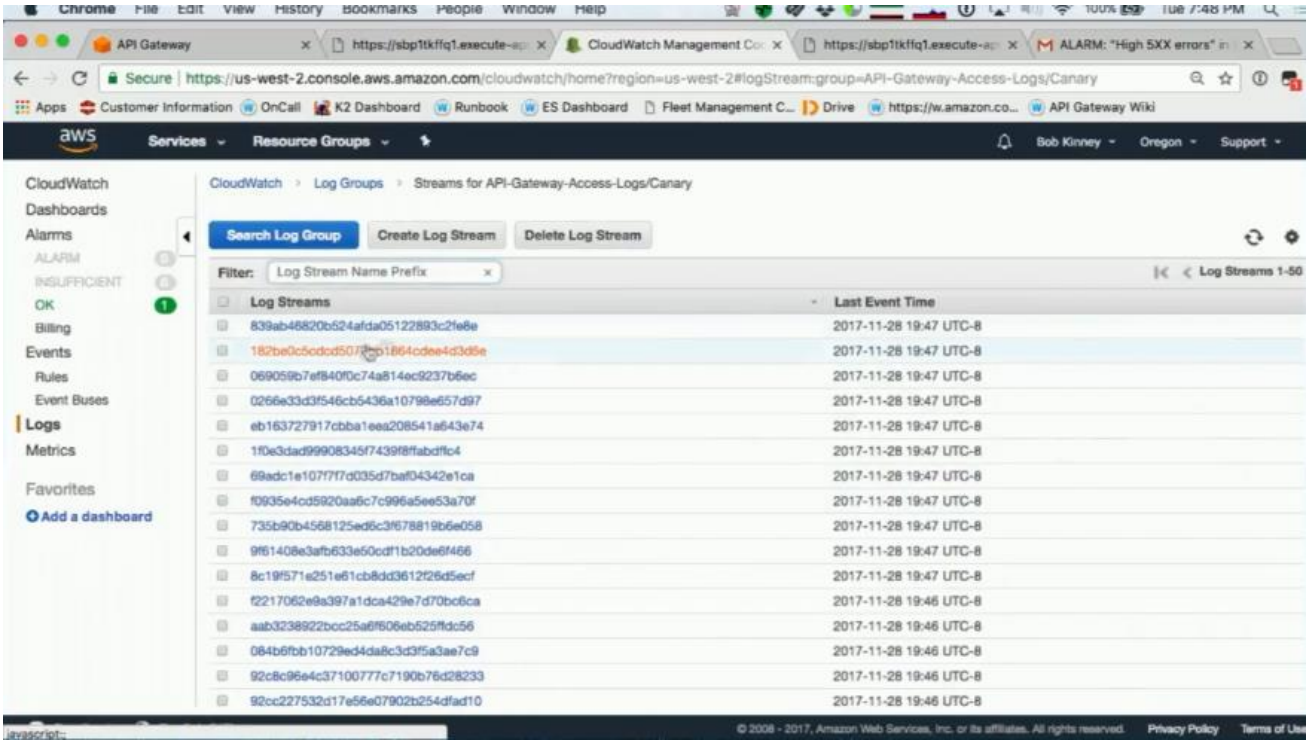


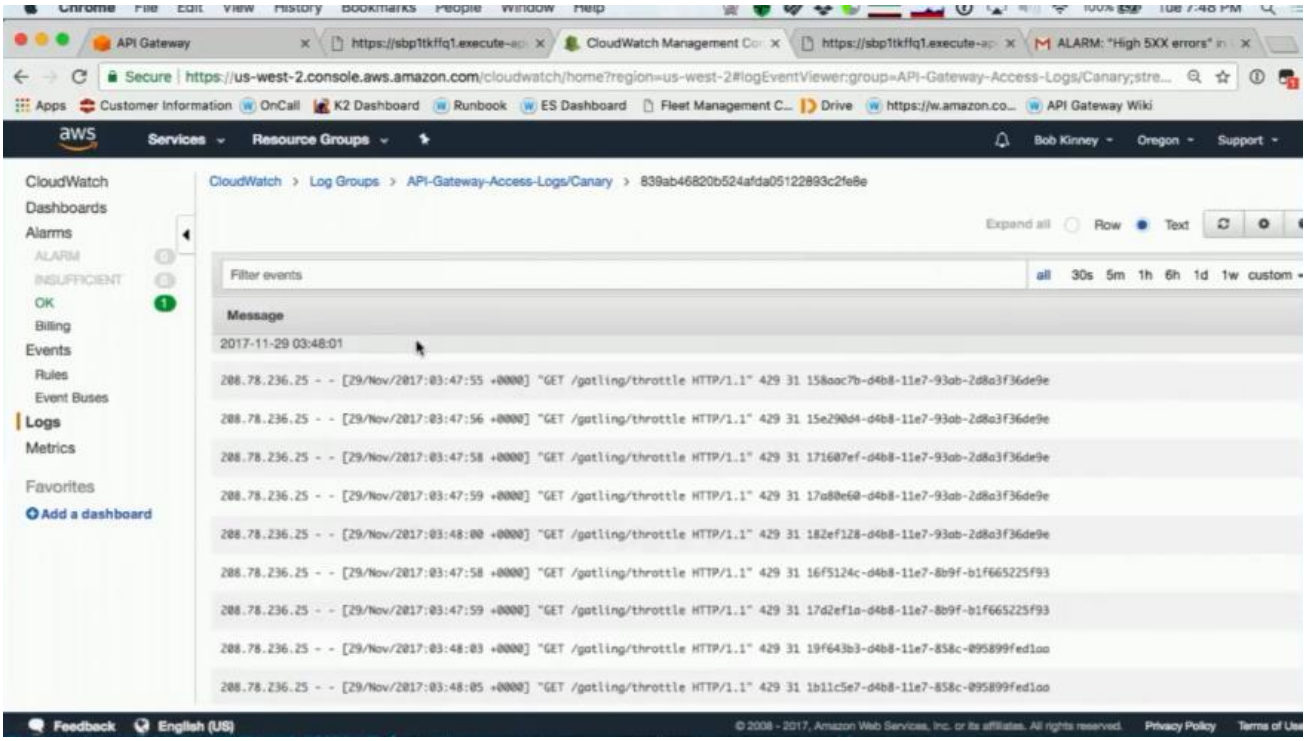
Note that we now have metrics for both the prod stage and the prod/Canary stage.



We should start seeing about 10% of the requests going to our canary API as above. This allows us to check the metrics for both stages and also set up alarms and maybe do automatic rollbacks, or automate the promotion of the canary to prod using a lambda

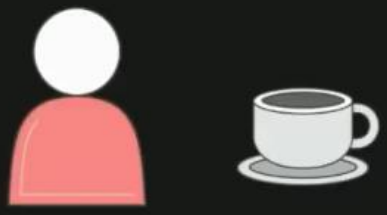






We can also see the logs for both the canary and prod stages

Check in with Doug



TAMPR is continuing to grow and Doug is now bringing in people to help work on updates.

He is looking for ways to formalize the update process.

API stages

Stages are named links to a deployed version of your API

Recommended for managing API lifecycle

- Dev/test/prod
- Alpha/beta/gamma

Support for parameterized values through *stage variables*

Custom domains

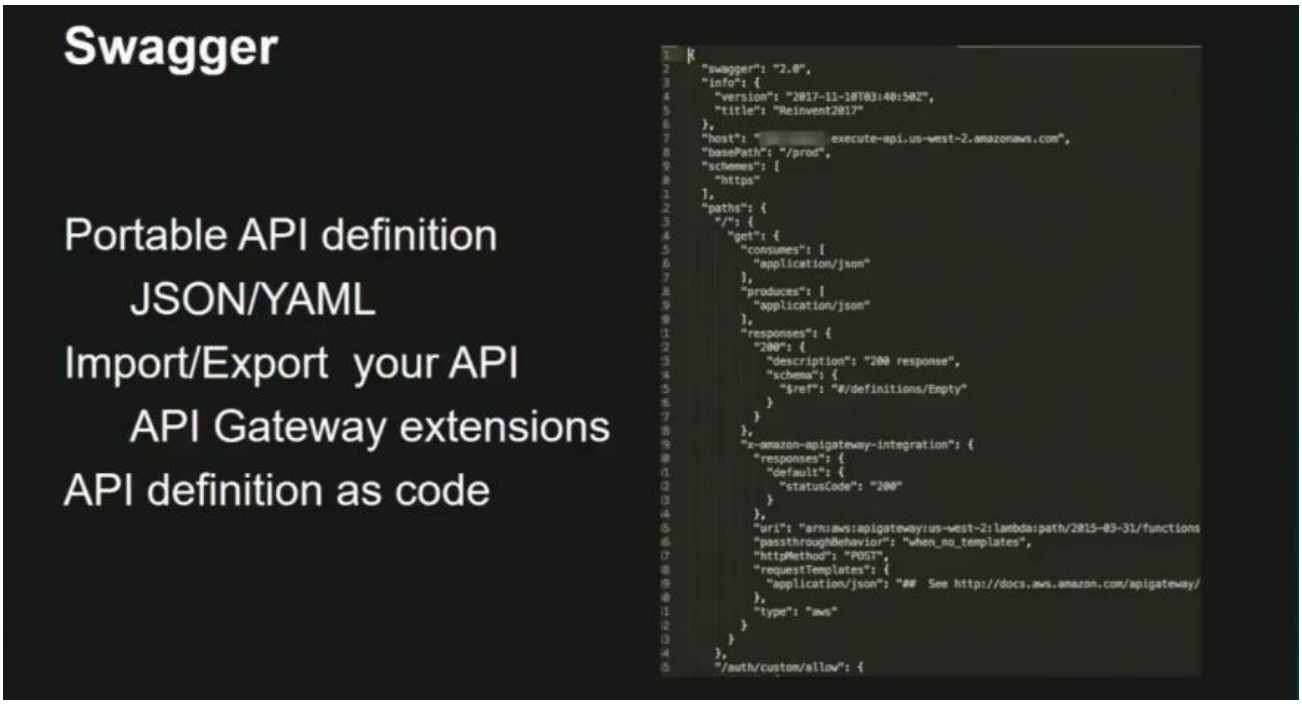
Run your APIs within your own DNS zone

Recommended for supporting multiple versions

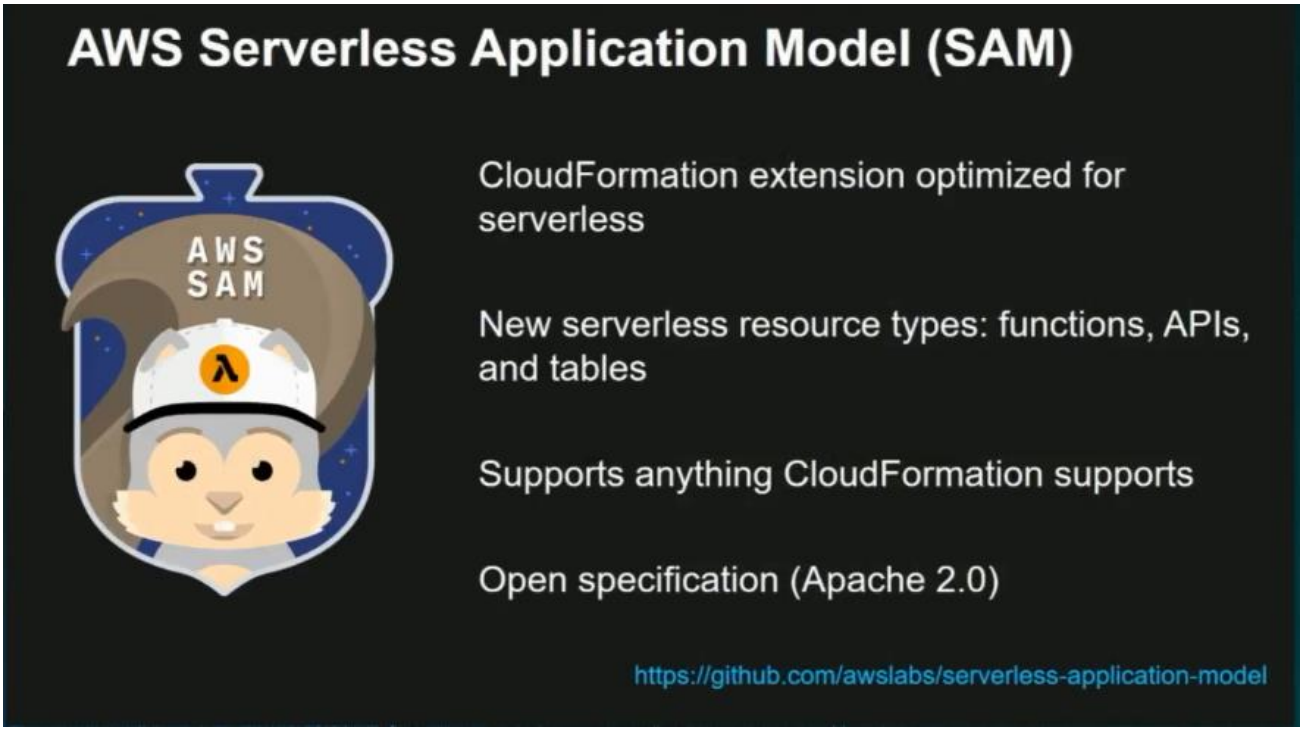
- api.tampr.com/v1 -> restapi1
- api.tampr.com/v2 -> restapi2

★ NEW★

Support for cross-region redundancy with regional API endpoints



We need a way to describe our APIs using a definition file that we can treat as code. The default way with API Gateway is using Swagger 2.0. we can get the swagger representation of our API if we want.



SAM template capabilities

- Can mix in other non-SAM CloudFormation resources in the same template
  - S3, Kinesis, Step Functions
- Supports use of Parameters, Mappings, Outputs, others
- Supports Intrinsic Functions
- Can use ImportValue
  (exceptions for RestApiId, Policies, StageName attributes)
- YAML or JSON

Other frameworks

APEX

aws Chalice

Claudia.js

SERVERLESS

SPARTA

Zappa

AWS CodeBuild

Fully managed build service that compiles source code, runs tests, and produces software packages

Scales continuously and processes multiple builds concurrently

You can provide custom build environments suited to your needs with Docker images

Only pay by the minute for the compute resources you use

Launched with CodePipeline and Jenkins integration

Can be used as a “Test” action in CodePipeline

AWS CodePipeline

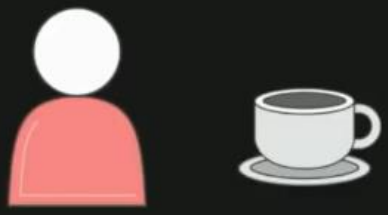
Continuous delivery service for fast and reliable application updates

Model and visualize your software release process

Builds, tests, and deploys your code every time there is a code change

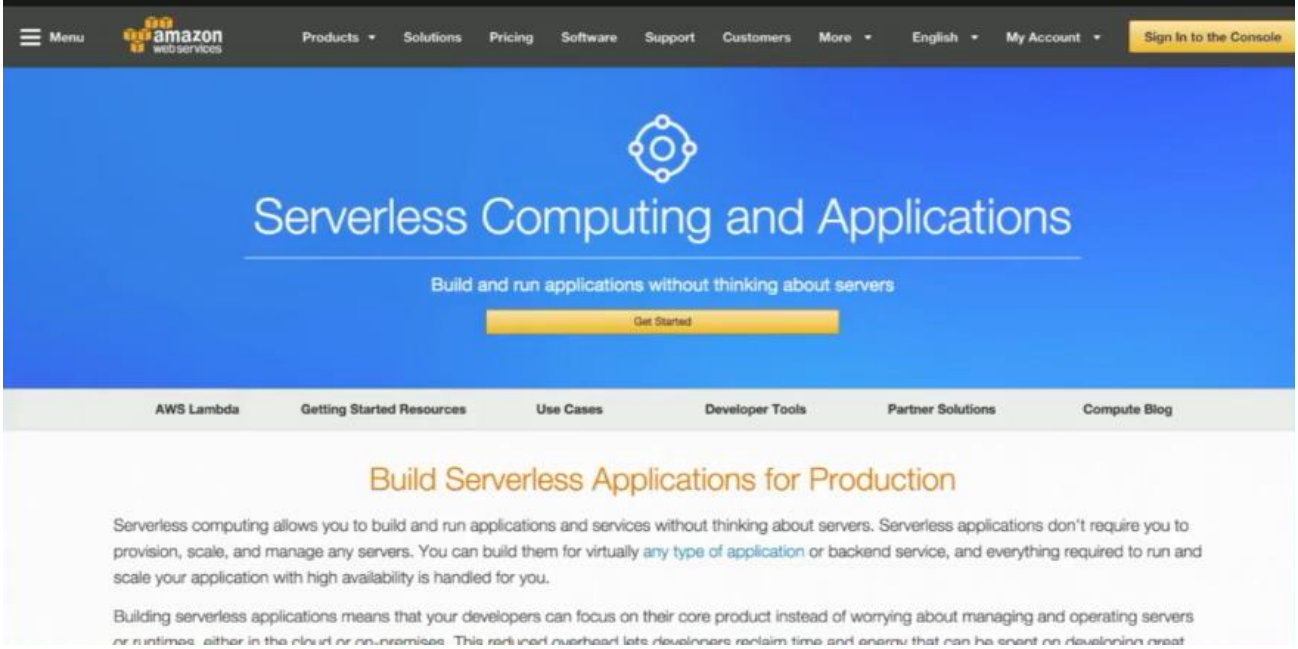
Integrates with third-party tools and AWS

Be like Doug



- Customize your API to meet your operational needs
- Monitor your APIs with metrics and alarms to find problems
- Use logging to diagnose problems with your APIs
- Make use of throttling and authentication to limit blast radius and protect critical API components
- Make your API available to third parties through usage plans
- Manage your API with the tools that make your development/deployment easier

aws.amazon.com/serverless



Other ways to connect

AWS Forums
- forums.aws.amazon.com

StackOverflow

GitHub
- github.com/aws
- github.com/awslabs