Dynamodb Demo

yupeng

05/23/14

Contents

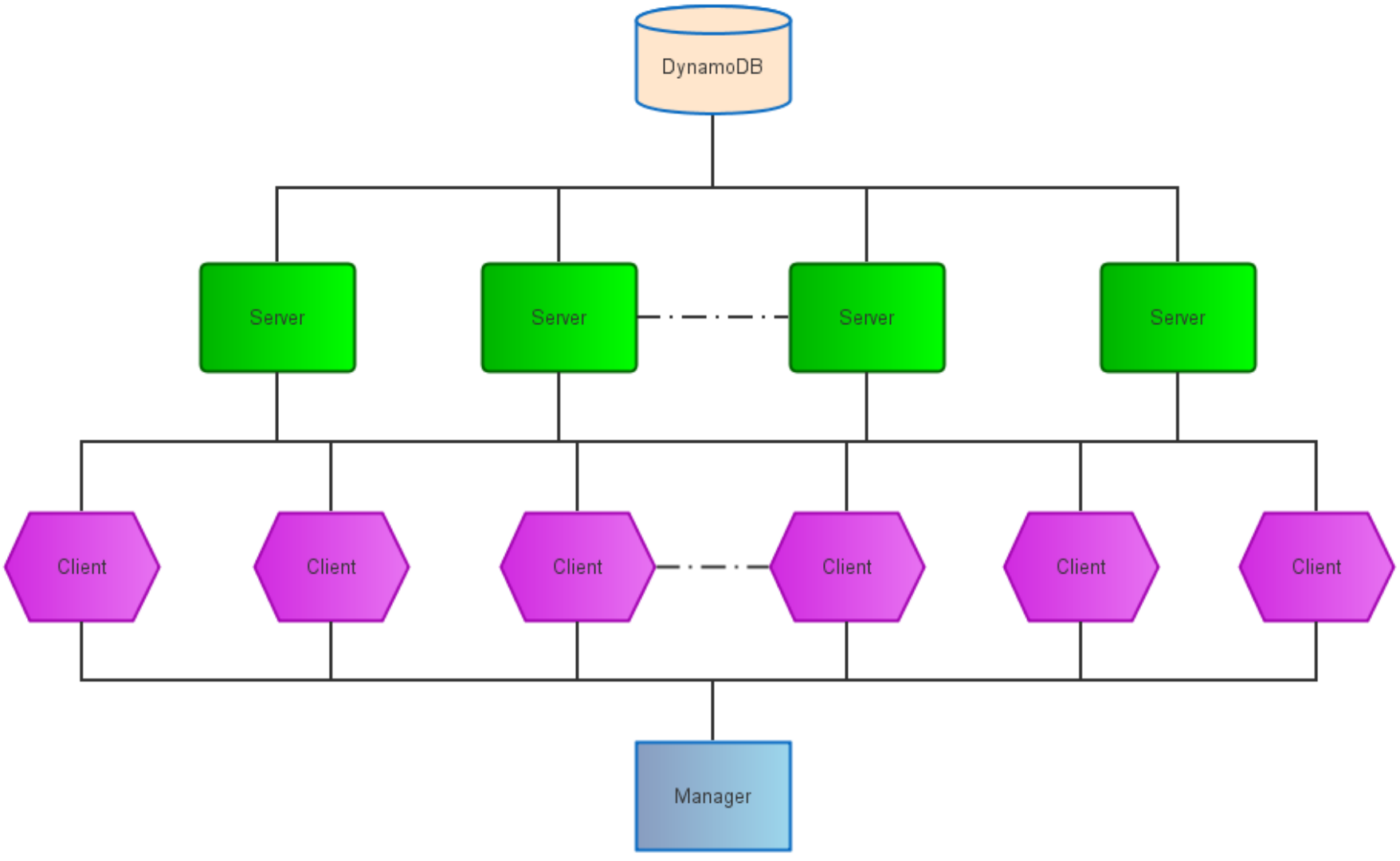
Main Sections

Architecure summary

The demo has 4 components.

- A dynamodb, store 10M lines data. The data schema is “name(string) date(string) score(int)”. It simulates a gaming database. Every user has a name, and many gaming scores ranged by date.
- Servers, simulate a gaming web server, provide a web interface to access the dynamodb
- Clients, using http_load generate workload to servers, http_load is an opensource web benchmark tool, we use it examine the servers’ respond latency.
- Manager, provide an interface to user, send command to clients, let clients generate workload by using http_load, and get the http_load result.

Architecture Picture



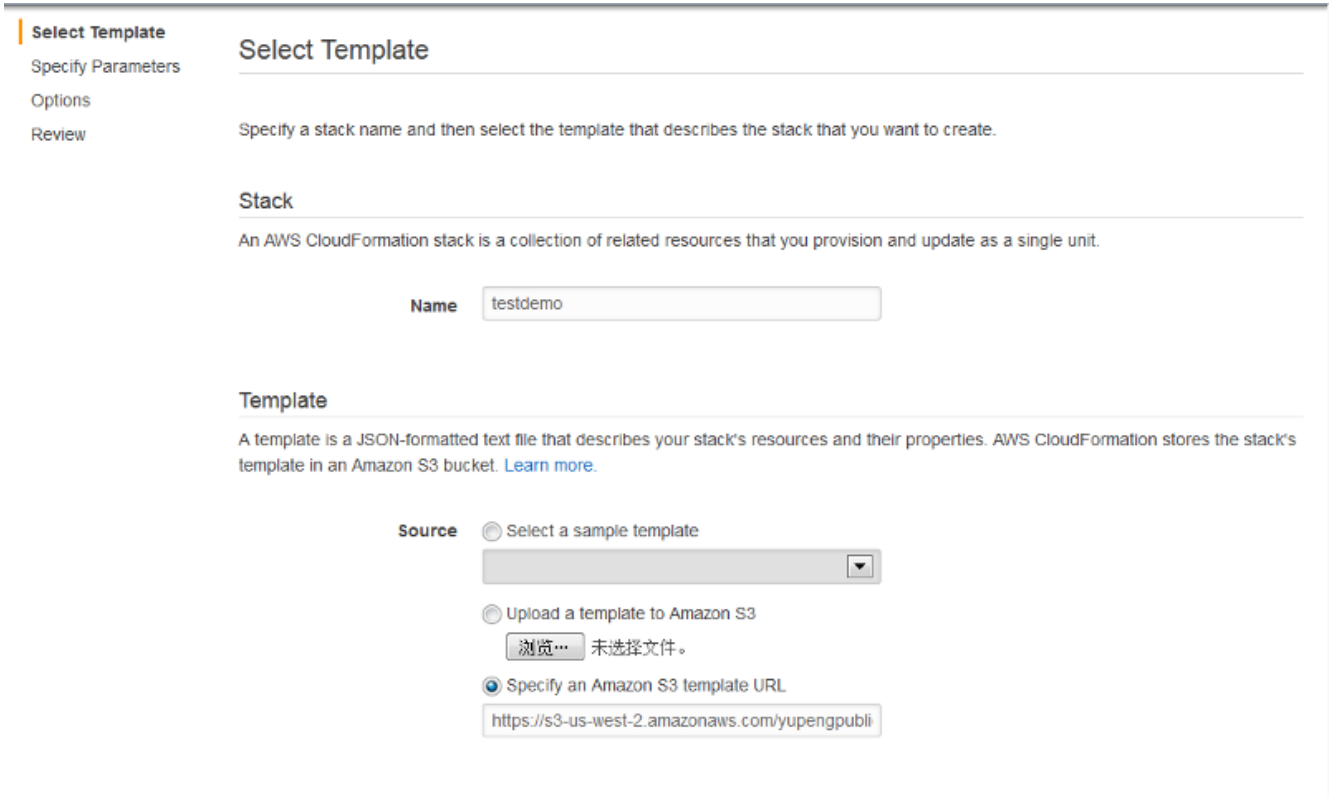
Main Sections

Cloudfromation template

The demo envornment can be deploied from a cloudformation template, the template address is: dynamodb_demo_link

Launch template

Launch the template in the cloudformation console:

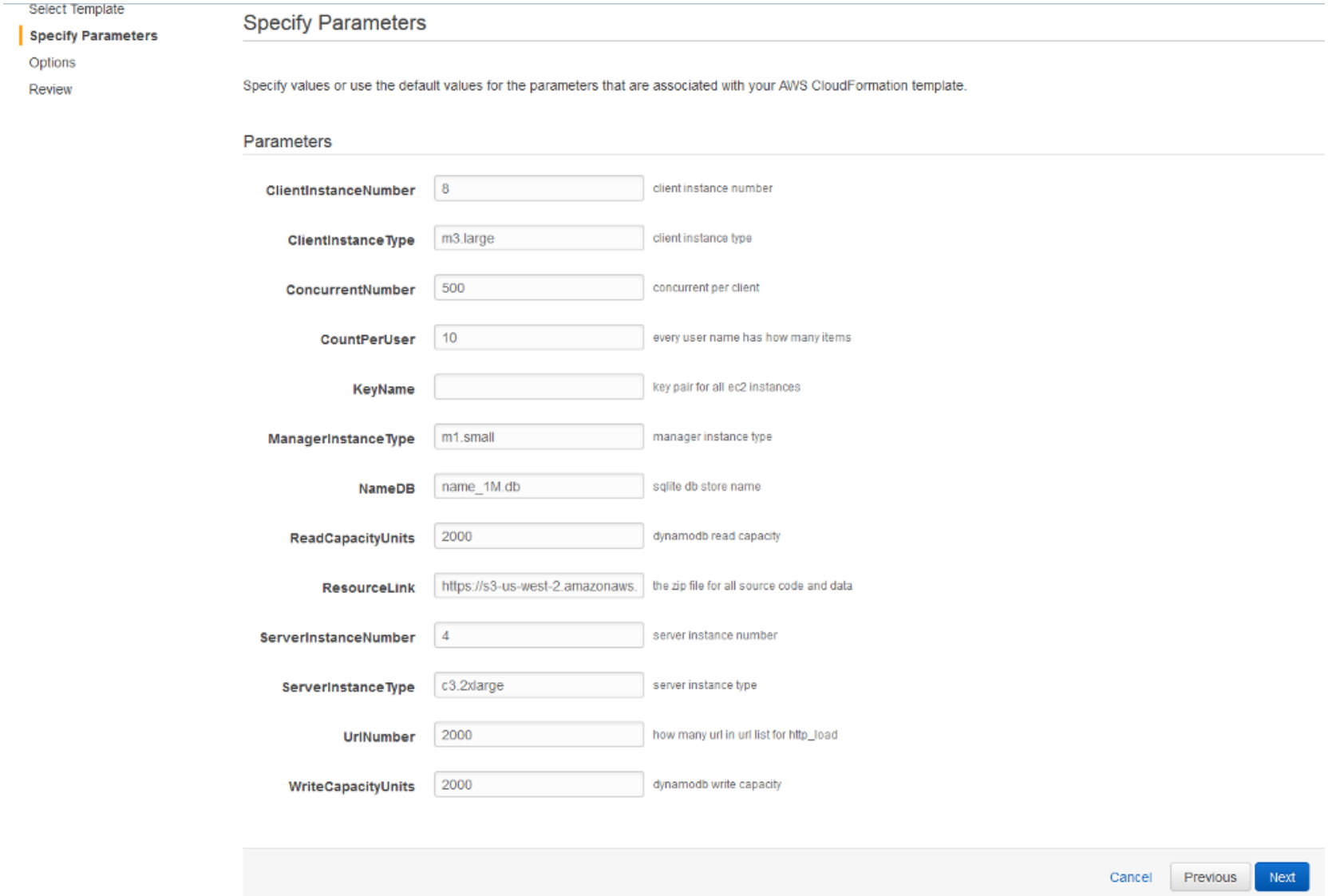


Template parameters

The template has several parameters. Most of them have default values. Only several parameters need to be set:

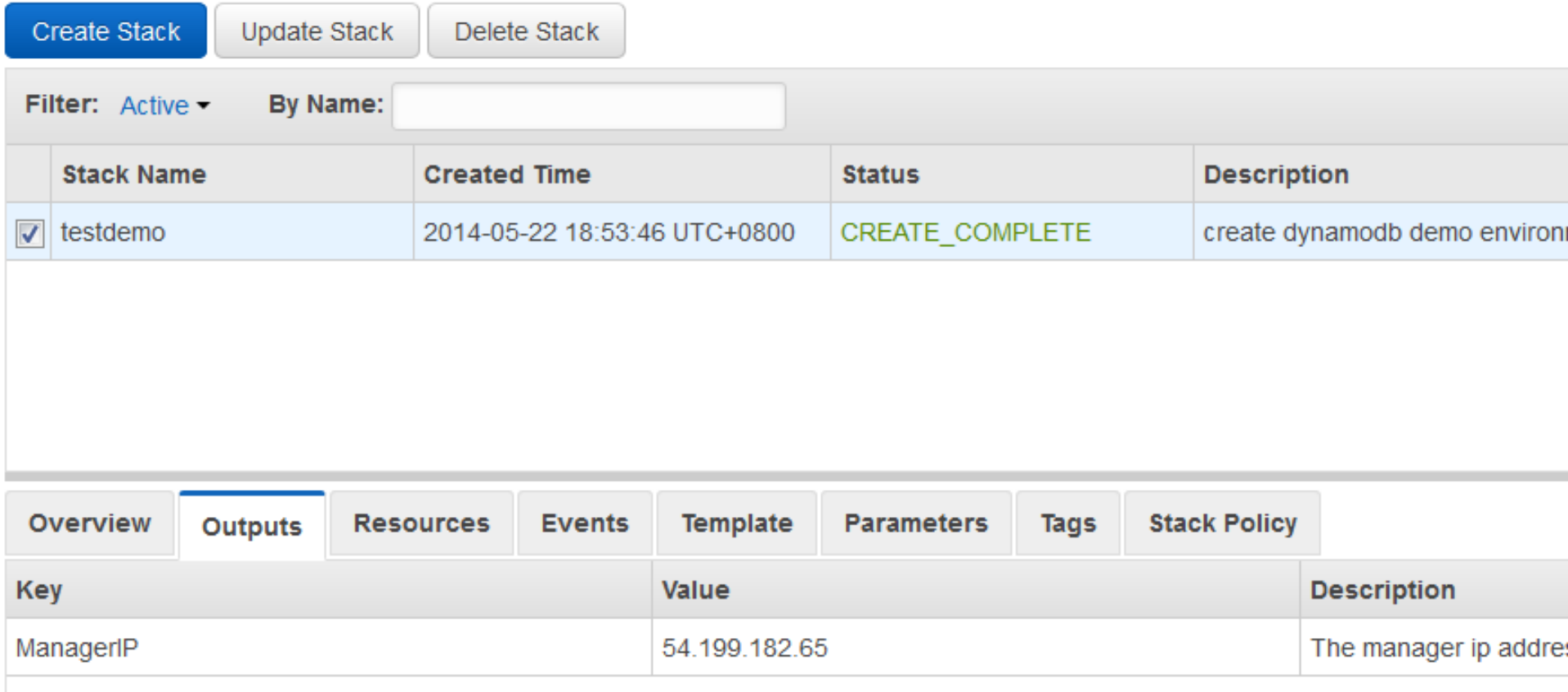
- KeyName, the keypair used for launch ec2 instance
- ReadCapacityUnits and WriteCapacityUnits, the dynamodb read/write capacity
- ServerInstanceNumber and ClientInstanceNumber, if the ReadCapacityUnits and WriteCapacityUnits are 2000, set ServerInstanceNumber to 4, set ClientInstanceNumber to 8, if the ReadCapacityUnits and WriteCapacityUnits are 4000, set ServerInstanceNumber to 8, set ClientInstanceNumber to 16

Parameter graph



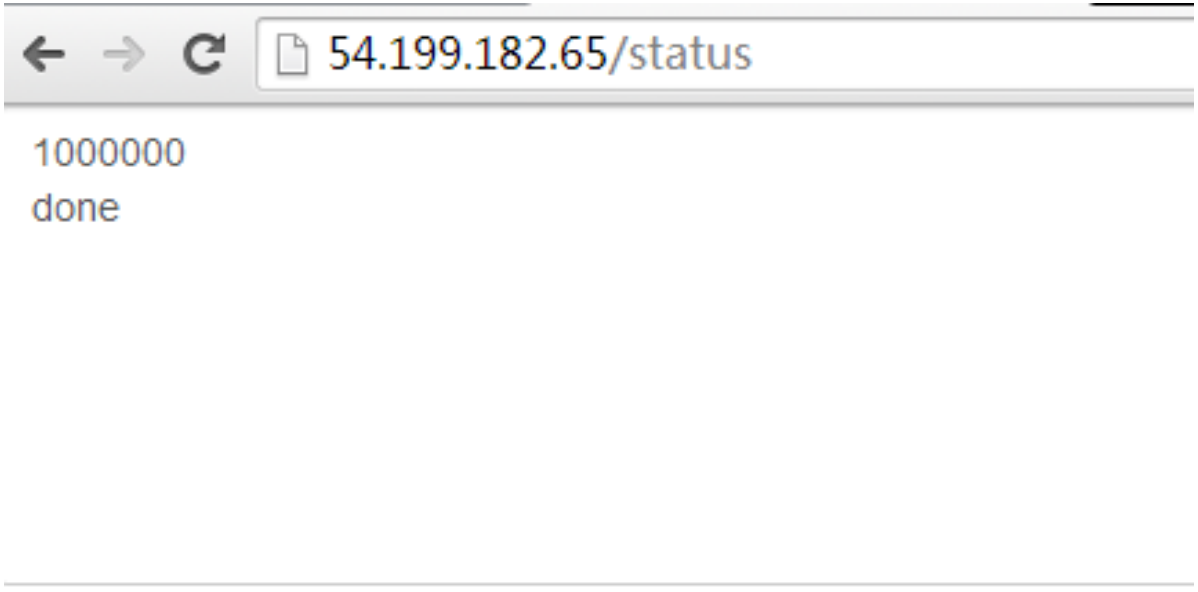
Check cloudformation Outputs

After the cloudformation launched complete, check the manager instance IP address in its Outputs



Check http://ManagerIP/status

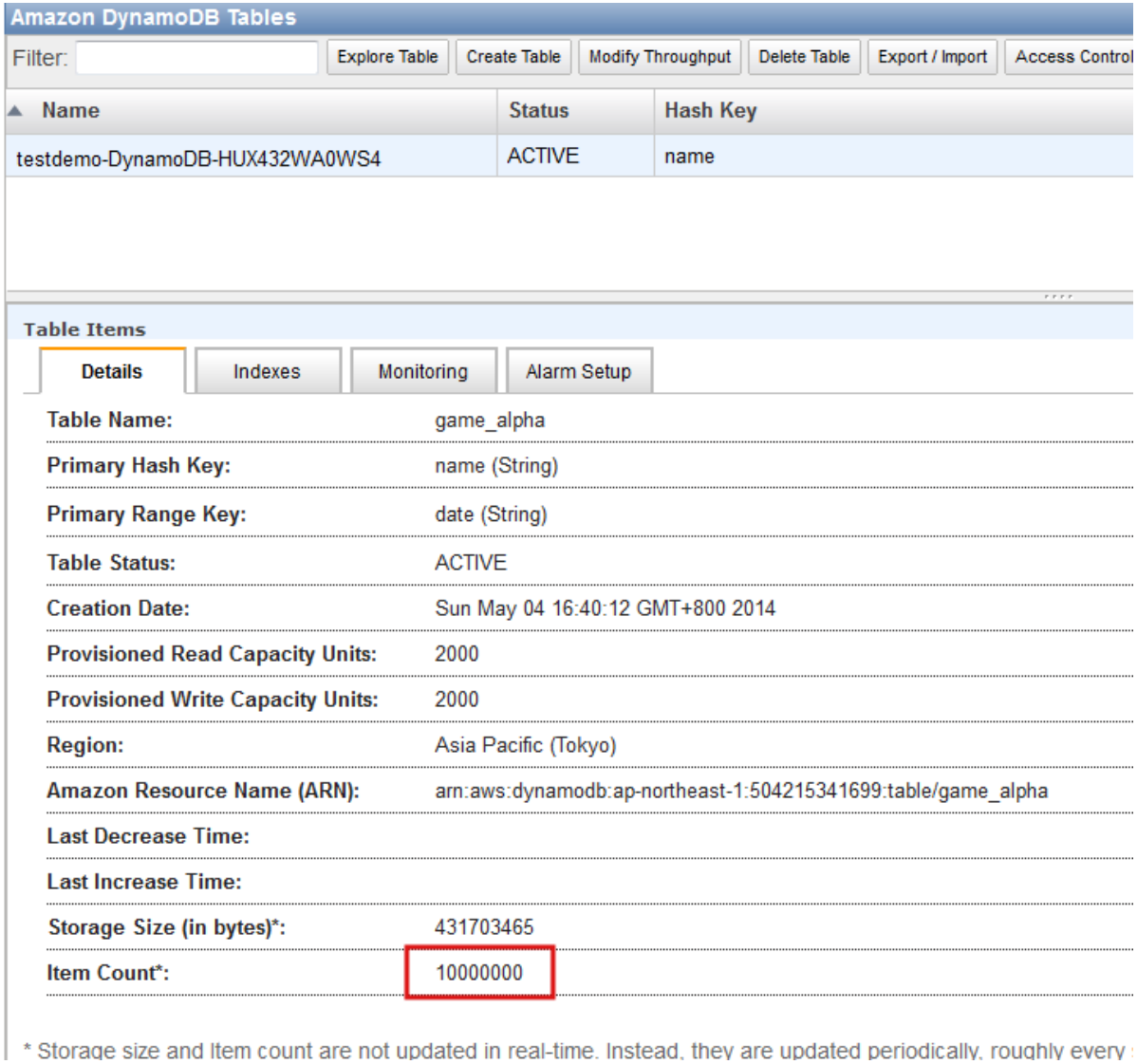
After the cloudformation launched, the manager instance will start to insert data to dynamodb, the http://ManagerIP/status indicate how many items are inserted. Seeing 'done' in that page indicate the insert complete, and we can start to demo.



Main Sections

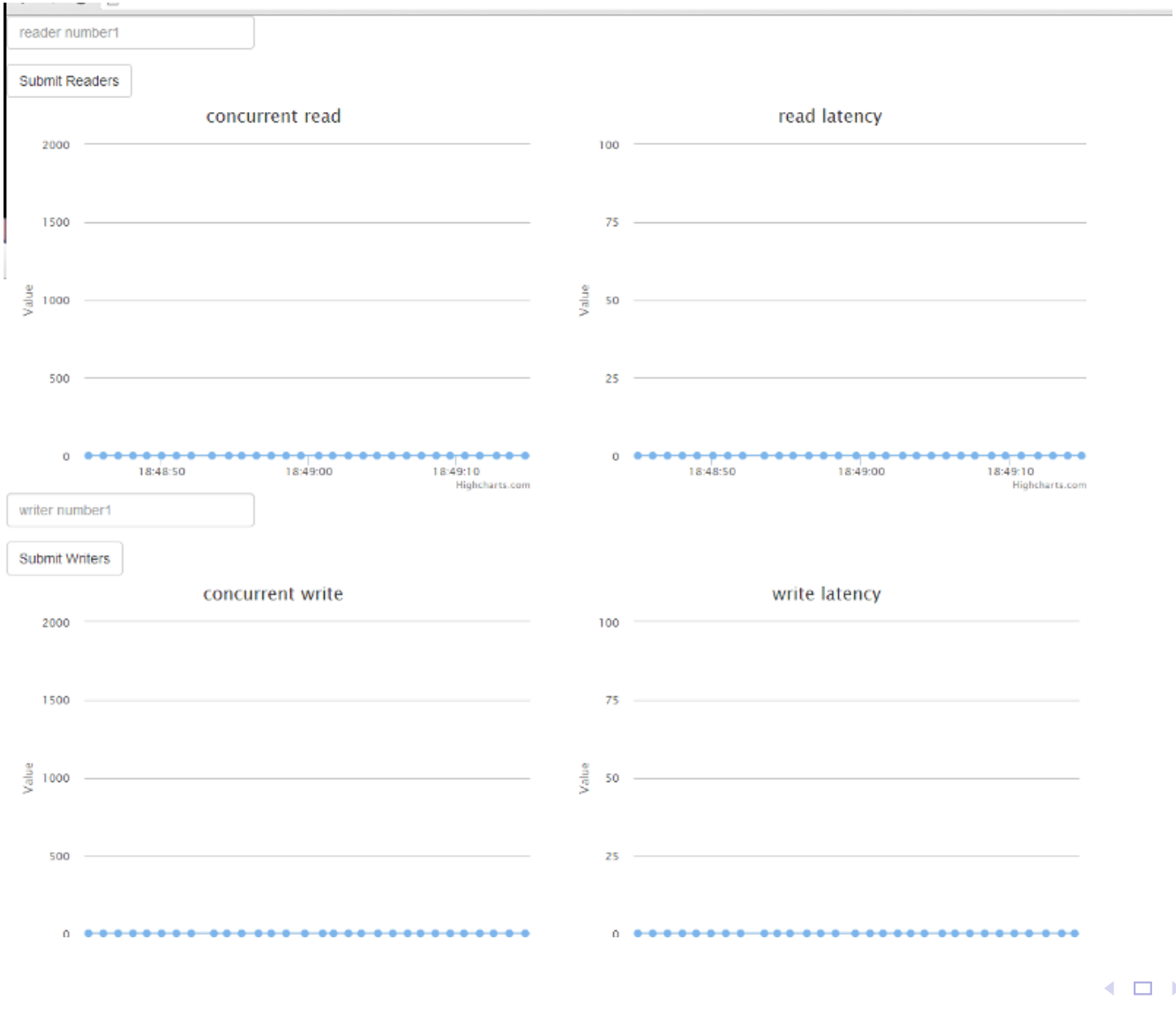
Table information

See the table information in the dynamodb console, examine the item count:



Manager Interface

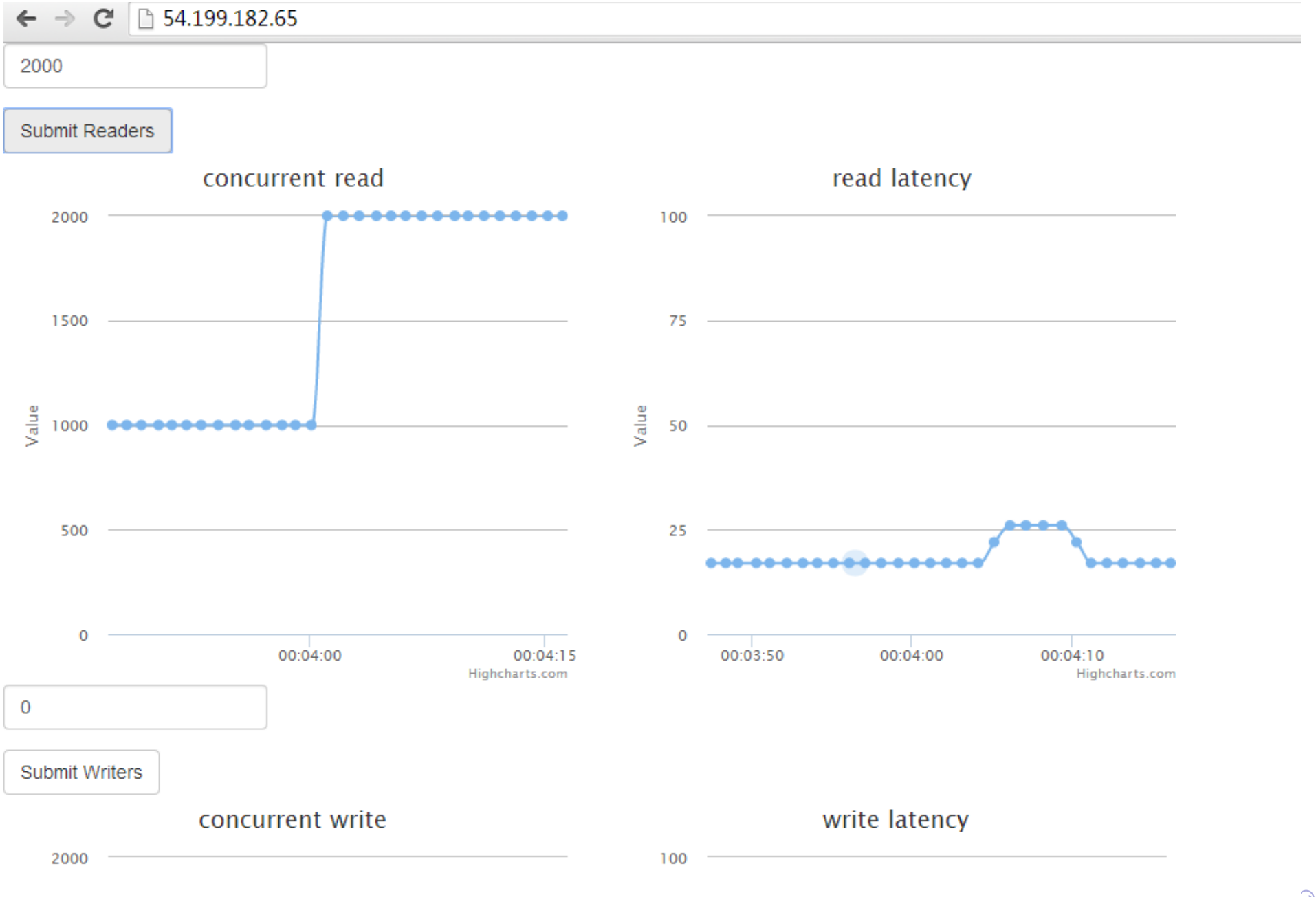
Input the manager instance's public in the web browser, you can get such an interface:



Generate read workload

Input 500, 1000, 1500, 2000 to the first form, and type “Submit Readers” can genreate 500, 1000, 1500, 2000 concurrent read to the dynamodb. The smallest grant size is 500, you can’t set to other value. The graph in the left top cornor show the concurrent read amount, and the graph in the right top cornor show the average read latency from the client perspective. Input 0 and submit the form will stop the workload.

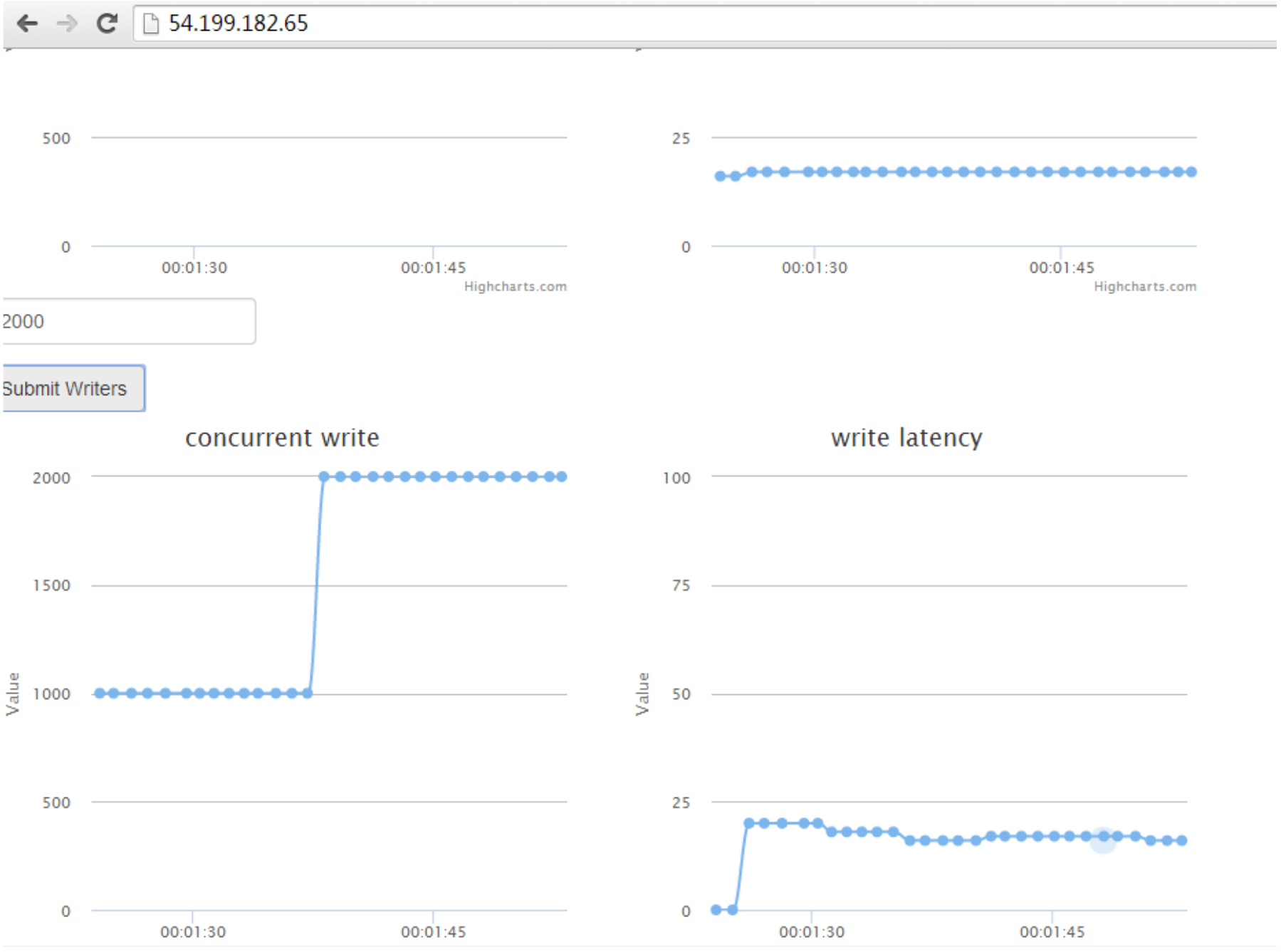
Read workload screenshot



Generate write workload

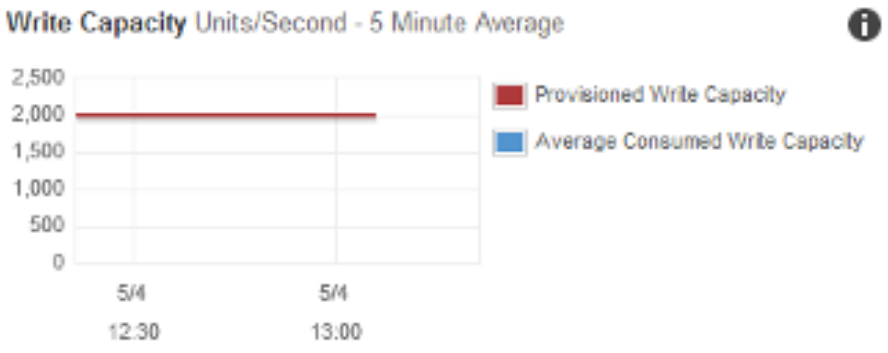
Input 500, 1000, 1500, 2000 to the second form, and type “Submit Writers” can generate 500, 1000, 1500, 2000 concurrent write to the dynamodb. The smallest grant size is 500, you can’t set to other value. The graph in the left bottom cornor show the concurrent write amount, and the graph in the right bottom cornor show the average write latency from the client perspective.

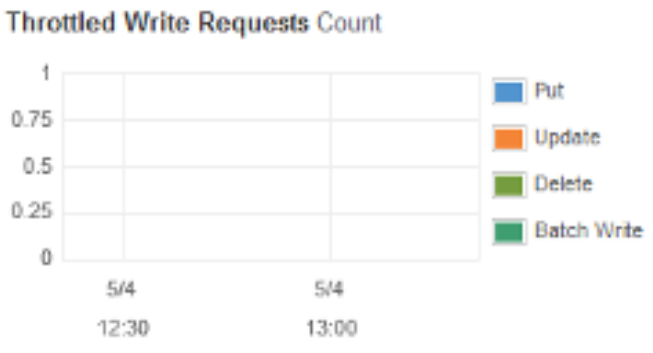
Write workload screenshot

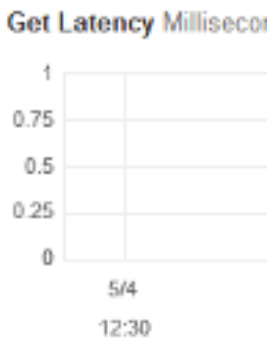


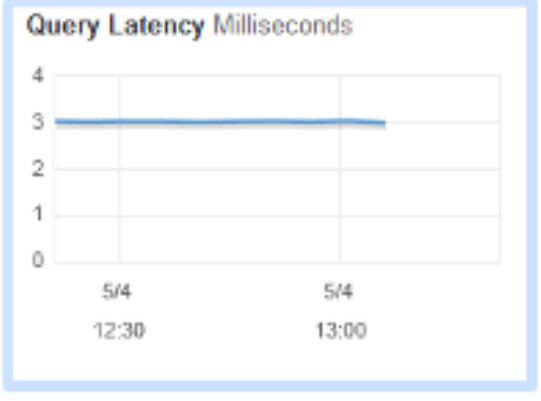
The dynamodb latency

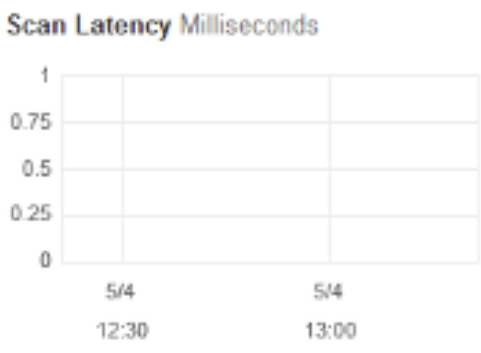
Check the dynamodb latency in cloudwatch





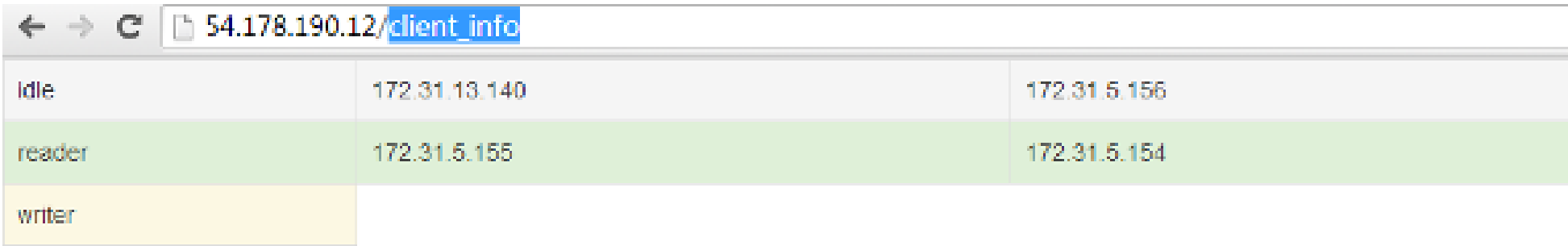






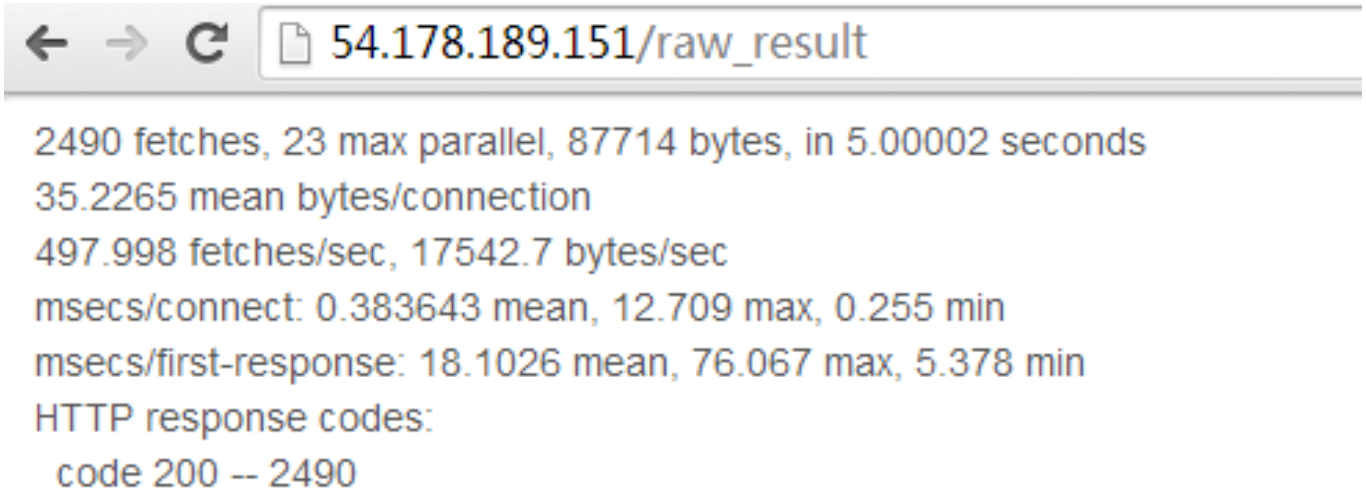
The http_load raw result

Check the server_ip/client_info page, we can see the ip address of the reader client, writer client and idle client



Check http_load raw result from client

From the ip address in server_ip/client_info page, we can choose a client, access the client_ip/raw_result page:



Main Sections

The demo soruce code

You can get the soruce code from github:
https://github.com/yupeng820921/dynamodb_demo.git
I also upload it to the S3:
https://s3-us-west-2.amazonaws.com/yupengpublic/dynamodb_demo.zip
The cloudformation has a parameter “ResourceLink”, it indicate where the server/client/manager instance should get the soruce code. You can modify the soruce code and upload to another place, and set the “ResouceLink” to another value.

Manage name

In the source code directory, we have a name_1M.db and name_20K.db. The name_1M.db have 1M random name, and name_20K.db have 20K random name. By default, manager will use name_1M.db, upload 10 “name(string) date(string) score(int)” items per name. If you want to change the name count, you can use the generate_name.py generate a new name db, store it to the source code directory, upload the new soruce code directory to S3, specific the filename of the “NameDB” parameter in cloudformation template. By default, the manager will upload 10 items per name, if you want to change it, you can assin a new value to “CountPerUser” parameter.

http_load parameters

We call http_load by such parameters:

http_load -p concurrent_number -r concurrent_number -s interval url_list.txt

Th

ConcurrentNumber and Interval parameters in the template wil be pass to the above command, and the UrlNumber parameter will be used to specific how many urls will be placed in a url_list.txt file.