Student ID: 452666 Name: Shuo Wu
Student ID: 451802 Name: Hanming Li

CSE 503 Performance Evaluation

1. Experiments choose:

- Apache web server performance evaluation on different AWS instance types (4 types)
- MongoDB performance evaluation on different instance types (4 types)

2. Experimental set up:

- Tool for apache web server analyze: apache bench (ab)
- Tool for MongoDB performance evaluation: Mongo-perf

We will run our experiments on the Amazon EC2 instances, using 4 different kinds of instances with different vCPU, ECU and memory. We will use different tools as mentioned above to get the data that we will use later for evaluation. The two experiments are contrast experiment that we just change specified factors such as vCPU or memory, and then show the impact of them to the performance of the web server (ec2 instances). What's more, after analyzing the feedback and comparing their prices, we will determine the highest cost-effective.

The following table is the information of different instances we will use:

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
General Purpose - Current Generation					
t2.nano	1	Variable	0.5	EBS Only	$0.0058 per Hour
t2.micro	1	Variable	1	EBS Only	$0.0116 per Hour
t2.small	1	Variable	2	EBS Only	$0.023 per Hour
t2.medium	2	Variable	4	EBS Only	$0.0464 per Hour
t2.large	2	Variable	8	EBS Only	$0.0928 per Hour
t2.xlarge	4	Variable	16	EBS Only	$0.1856 per Hour
t2.2xlarge	8	Variable	32	EBS Only	$0.3712 per Hour
m4.large	2	6.5	8	EBS Only	$0.1 per Hour
m4.xlarge	4	13	16	EBS Only	$0.2 per Hour
m4.2xlarge	8	26	32	EBS Only	$0.4 per Hour
m4.4xlarge	16	53.5	64	EBS Only	$0.8 per Hour

(we will use t2.nano, t2.micro, t2.medium and m4.large to analyze)

By choosing these instances, we can get the impact of memory by comparing t2.nano and t2.micro, we can get the impact of vCPU by comparing t2.micro and t2,medium and ect.

Student ID: 452666 Name: Shuo Wu
Student ID: 451802 Name: Hanming Li

3. Apache web server performance evaluation

The command line we use to run the apache bench on instances:

wushuo$ ab -c 10 -n 1000 http://13.58.151.103/

wushuo$ ab -c 100 -n 1000 http://13.58.151.103/

Using ab to send 1000 (-n change the total number of requests) request, 10 or 100 (-c change the number of concurrency) at a time to the URL 13.58.151.103.

One of the result:

```
Server Port:            80

Document Path:          /
Document Length:        4891 bytes

Concurrency Level:      100
Time taken for tests:   0.400 seconds
Complete requests:      1000
Failed requests:        0
Non-2xx responses:      1000
Total transferred:      5174000 bytes
HTML transferred:       4891000 bytes
Requests per second:    2503.11 [#/sec] (mean)
Time per request:       39.950 [ms] (mean)
Time per request:       0.400 [ms] (mean, across all concurrent requests)
Transfer rate:          12647.55 [Kbytes/sec] received

Connection Times (ms)
              min  mean[+/-sd] median   max
Connect:        2    4   1.5      3       8
Processing:    13   35   5.0     36      45
Waiting:        7   34   5.0     35      43
Total:         15   39   4.9     39      53
```

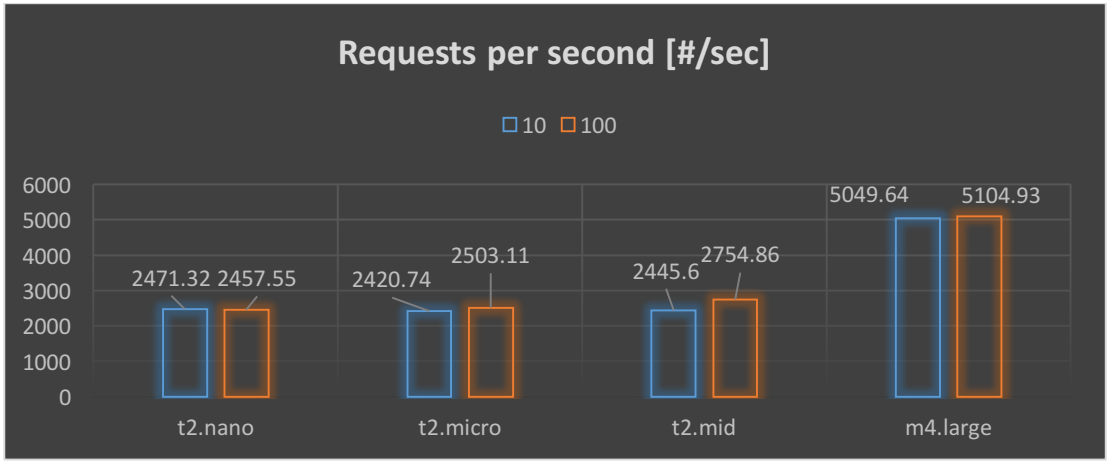

Figure 1: Apache requests per second

As mentioned before, we also get some data related to connection times. For example, we get connect time, processing time, waiting time and total time. We believe that the number of request per time is an important factor of processing efficiency. Therefore, we plot the output data base on 10 requests per time and 100 requests per time.

Student ID: 452666
Name: Shuo Wu
Student ID: 451802
Name: Hanming Li

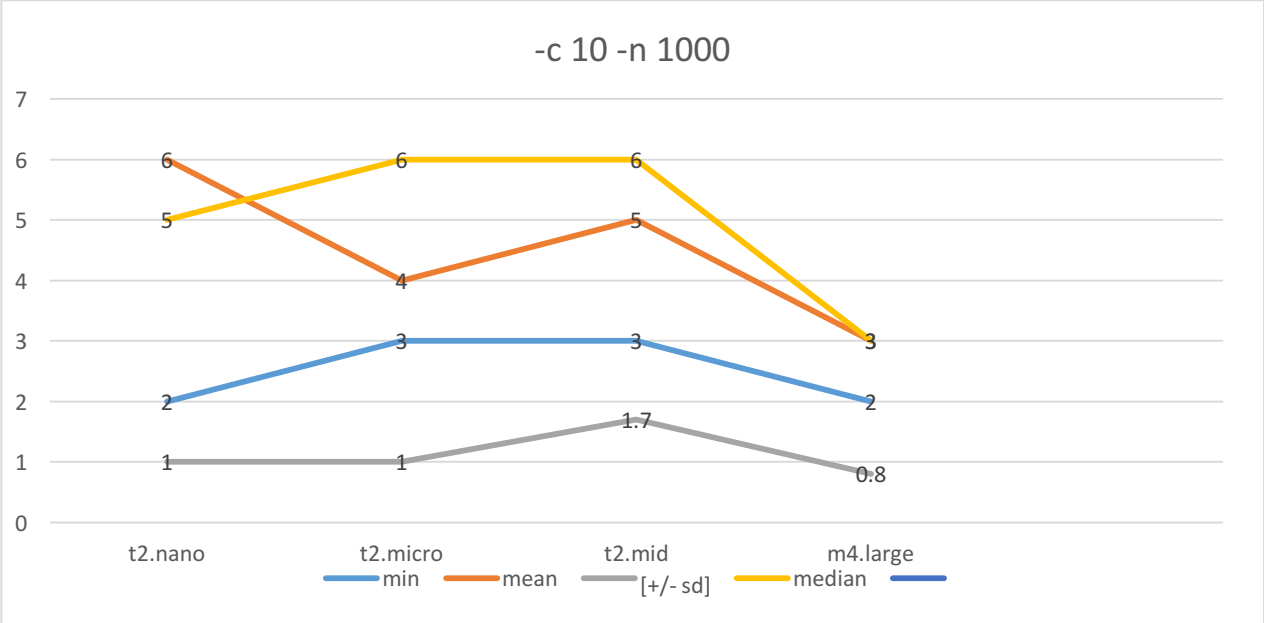

Figure 2 (A): Apache connection time

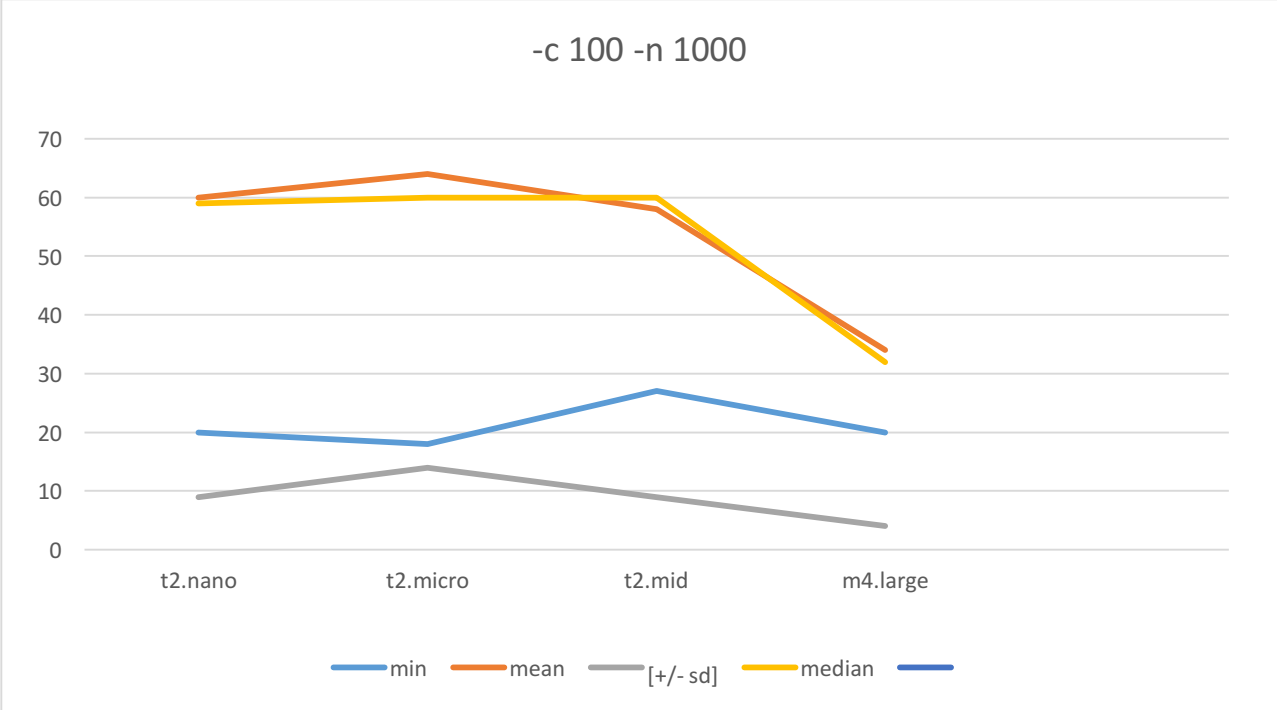

Figure 2 (A): Apache connection time

From the graph, we can see a huge gap when the RAM of instance increase. For example, we can see the m4.large with a small connection time when dealing the requests. Also, we can tell from the result that the vCPU doesn't play a very important role in Apache web performance. Therefore, we conclusion that Memory of instance is more important than vCPU in web performance.

Student ID: 452666 Name: Shuo Wu
Student ID: 451802 Name: Hanming Li

4. MongoDB Performance Evaluation

As I mentioned above, we will use the Mongo-perf to evaluate the MongoDB performance. Our team will test the writing process and reading process from the database. Since these two operations are very common and basic operations of a database, and could be very representative. Therefore, the following part will use the data related reading and writing part.

We would only use the empty insertion and empty query because the result would be affected by the different size or content of the databases, and could represent the overall performance of the empty database. Another reason is that some of the test set up seems may case the database to crash and make the benchmark terminated with error. To avoid it, we decided to do the smallest scale test on the web instances. We used different threads : 1, 2, 4, 8 and 16 to test the performance of database under different pressure.

The following graph are the bench result:

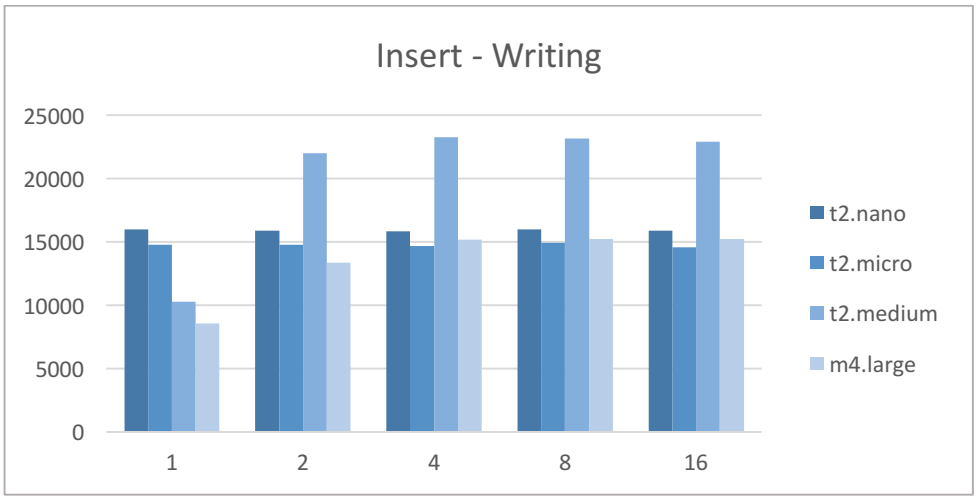

Figure 3(A): Insert of MongoDB (Writing)

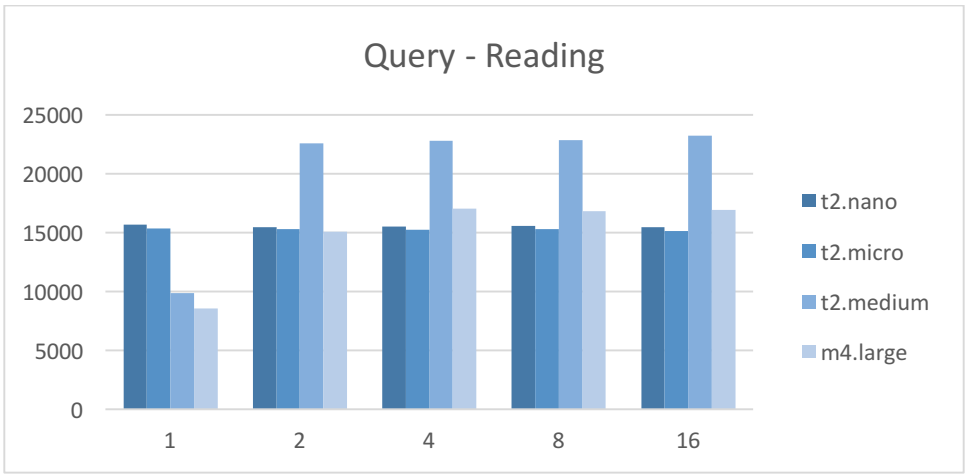

Figure 3(B): Query of MongoDB (Reading)

From the data we get, we can say that vCPU plays an important role in database performance. While the memory might be not.

For example, comparing the outcome of t2.micro (with 1 vCPU & 4 GiB RAM), in order to make the same memory, we just changed the memory size or t2.micro instance and t2.medium (with 2 vCPU & 4 GiB RAM). Although the one thread performance is kind of the same(the first one is better than the second one), the performance boost by RAM is about 30% on both single thread on multi-thread. It is reasonable since one more core means two thread can insert into the database simultaneously which result in better performance.

While, when comparing the data related to t2.nano(with 1 vCPU & 0.5 GiB RAM) and t2.micro (with 1 vCPU & 1 GiB RAM), there isn't a signification change between these two instances. Therefore, we could say memory doesn't impact the database performance a lot. However, that might be the thing need us attention. This may be caused by the CPU bottleneck effect as the only one core is unable to fully utilize the memory. What's more, we can't deny that the different location of the machine that hold the instances may give some impact on the performance of database.

5. Discussion

Hardware properties will influence the performance of software. From the research above, we can say the better hardware (more CPU or core and more memory), the better performance of a server if we don't care about the price. For the bottleneck, we can say if a server has limited CPU, the size of memory doesn't make any sense, since the CPU can't handle the requests. If I were asked to recommend a instance, t2.medium would be the way to go. Comparing with its performance and price.