

CS 5296 Spring 2025 Assignment 1 (2 questions, 5 marks)

1. Measure EC2 CPU and Memory performance

| Size | CPU performance -- total time (s) | Memory performance --transfer speed MB/s |
|------------------|--------------------------------------|---|
| t2. micro | 10.0006s | 536.94 MB/sec |
| t2.medium | 10.0008s | 900.67 MB/sec |
| t2.large | 10.0009s | 941.87 MB/sec |

0.9 total time should be reducing
CPU performance measurement command:

“sysbench cpu --threads=4 --cpu-max-prime=10000 run”

Memory performance measurement command:

“sysbench memory --threads=4 --memory-total-size=10G --memory-oper=write --memory-scope=global run”

Measurement Analysis:

This analysis evaluates the CPU and memory performance of EC2 t2.micro, t2.medium, and t2.large instances using sysbench.

CPU Performance: Execution time remains around 10s across all instances due to t2 burstable CPU limits. No significant improvement with larger instances.

Memory performance improves with instance size, but gains diminish at higher tiers.

2.Measure EC2 **network** performance

1) The network performance is experienced between instances of the same type and different types

| Type | TCP bandwidth(Mbps) | Average RTT (ms) |
|-----------------------|---------------------|------------------|
| t2.micro - t2.micro | 447 Mbps | 1.237 ms |
| t2.micro - t2.medium | 542 Mbps | 1.282 ms |
| t2.medium - t2.medium | 676 Mbps | 1.015 ms |
| t2.medium - t2.large | 818 Mbps | 0.973 ms |
| t2.large - t2.large | 910 Mbps | 0.642 ms |

2) The network performance (bandwidth) with different window sizes

Q2(2)-(5), 2 marks

| Window Size | TCP bandwidth (Mbps) |
|-------------|----------------------|
| 128K | 387Mbps |
| 256K | 465Mbps |
| 512K | 545Mbps |

For RTT, ping from different instances (clients) to the same instance (server) simultaneously

| Number of Clients | Client1 Average RTT (ms) | Client2 Average RTT (ms) | Client3 Average RTT (ms) |
|-------------------|-----------------------------|-----------------------------|-----------------------------|
| 2 | 1.245ms | 1.347ms | none |
| 3 | 1.412ms | 1.664ms | 1.842ms |

3) The network performance with one server and multiple clients

| Number of Clients | Client1 TCP bandwidth (Mbps) | Client2 TCP bandwidth (Mbps) | Client3 TCP bandwidth (Mbps) |
|-------------------|------------------------------|------------------------------|------------------------------|
| 2 | 556Mbps | 553Mbps | none |
| 3 | 413Mbps | 389Mbps | 376Mbps |

4) The network performance for instances deployed in different regions

| Type | TCP bandwidth (Mbps) | Average RTT (ms) |
|----------------------------|----------------------|------------------|
| N. Virginia - Oregon (I) | 15.1Mbps | 74.126ms |
| N. Virginia - Oregon (II) | 26.8Mbps | 68.482ms |
| Oregon - N. Virginia (III) | 36.7Mbps | 65.237ms |

5) The network performance in different time

| Time(HKT) | TCP bandwidth (Mbps) | Average RTT (ms) |
|---------------------|----------------------|------------------|
| Morning (~10:00am) | 812Mbps | 1.128ms |
| Afternoon (~4:00pm) | 758Mbps | 1.321ms |
| Evening (~10:00pm) | 672Mbps | 1.786ms |

6) Open-ended question: so far you have measured network performance in different scenarios. Observe the values in each table, and try to explain why network performance varies under different scenarios?

Q2(6), 0.5

- (1)** Larger instances, such as t2.large, have higher allocated network bandwidth compared to smaller instances like t2.micro. AWS assigns more network resources to larger instances, leading to better performance.
- (2)** Intra-region communication (within the same AWS region) has significantly lower latency (~1-2ms) and higher bandwidth. Inter-region communication requires routing through AWS backbone infrastructure, increasing latency (~65-75ms) and reducing throughput.
- (3)** When multiple clients communicate with a single server, available bandwidth is shared among them. More clients result in lower per-client bandwidth due to the fixed network resources allocated to the instance.
- (4)** Network performance fluctuates throughout the day. Mornings generally provide the best bandwidth and lowest latency, while afternoons see moderate network load. Evenings often experience the highest congestion, leading to reduced throughput and increased latency.
- (5)** Larger TCP window sizes improve bandwidth utilization by reducing the number of acknowledgments required, optimizing performance. Smaller window sizes limit throughput due to increased protocol overhead.