Cloud Computing first assignment 2020

Group 25

November 2020

participants
amro.hendawi
Shirkouh Matini
Ammer Ayach
Abdelkader Alkadour
Patrick Andreas Stadler

1 CPU benchmark questions

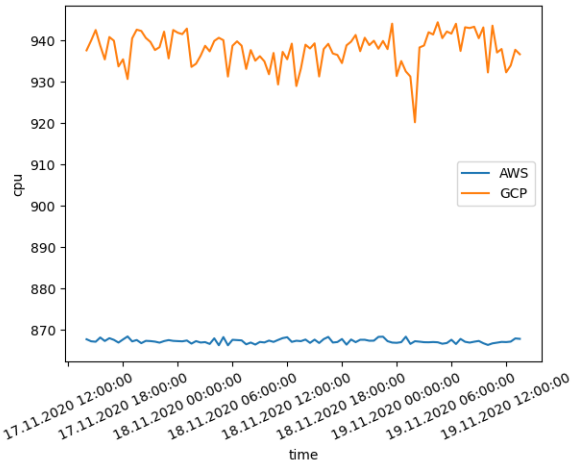

Figure 1: CPU Benchmark results

1. Shortly describe, how sysbench performs CPU benchmark. What does the resulting events/s value represent?

Sysbench evaluates CPU performance by verifying prime numbers by doing standard division of the number by all numbers between 2 and the square root of the number. If any number gives a remainder of 0, the next number is calculated putting some stress on the CPU, but only on a very limited set of the CPU's features. It is also possible to specify the number of threads running, making it a quick and easy way to test CPU performance for various numbers of threads.

2. Look at the plots of your long-term measurements. Do you see any seasonal changes?

   The CPU plot in figure 1 shows various fluctuations, especially in GCP benchmark line, yet no seasonal changes.

2 Memory benchmark questions

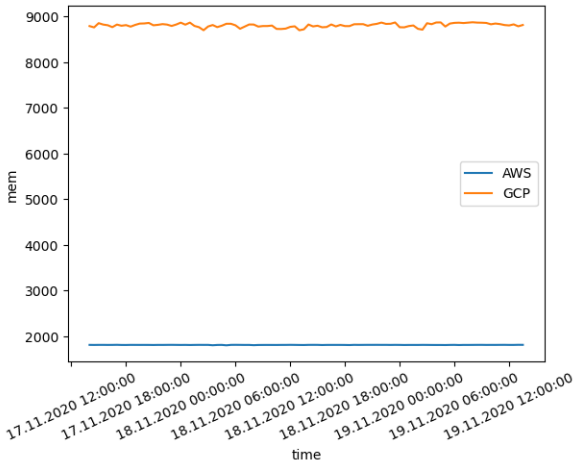

Figure 2: Memory benchmark results

1. Shortly describe, how sysbench measures memory performance.

   The benchmark application allocates a memory buffer in order to write and read from it, each time for the size of a pointer, whether it's 32bit or 64bit, and each execution until the total buffer size has been read from or written to. This is then repeated until the provided volume (–memory-total-size) is reached. Users can provide multiple threads (–num-threads), different sizes in buffer (–memory-block-size) and the type of requests

(read or write, sequential or random).
The program collects timestamps before and after the reading/writing process, which then illustrates through plots how good the memory performs under heavy load.

2. How would you expect virtualization to affect the memory benchmark? Why?

   Memory benchmarking measures the required time to write and clean heap memory from different locations. The locations are chosen such that a cache miss occurs and data is loaded directly from memory. Since the hypervisor needs to communicate at hardware level for memory requests we get considerably better results for memory in non-virtualized systems.

3 Disk benchmark questions

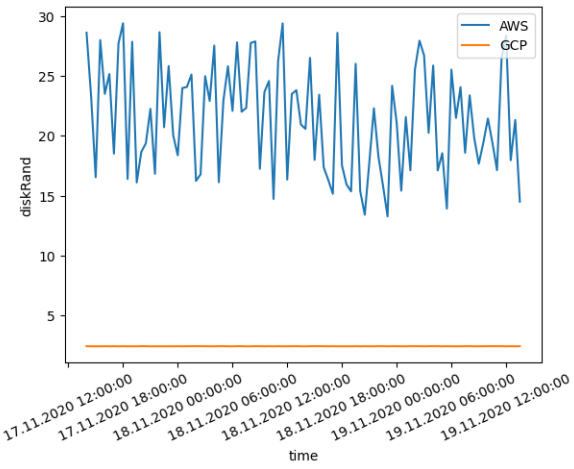

Figure 3: random disk read benchmark results

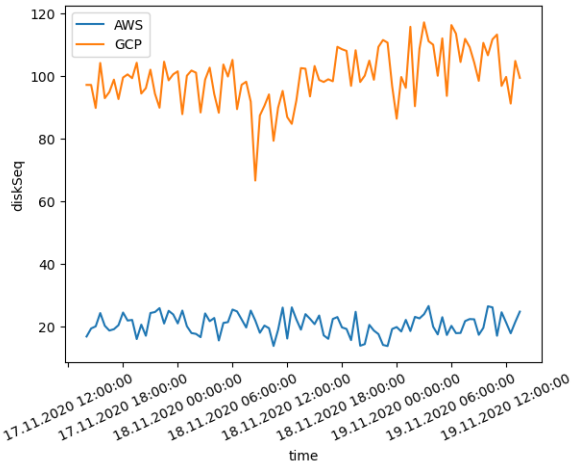

Figure 4: sequential disk read benchmark results

1. Shortly describe, how sysbench performs the disk benchmarks.

   At the prepare stage SysBench creates a specified number of files with a specified total size, then at the run stage, each thread performs specified I/O operations on this set of files. When the global –validate option is used with the fileio test mode, SysBench performs checksums validation on all data read from the disk. On each write operation the block is filled with random values, then the checksum is calculated and stored in the block along with the offset of this block within a file. On each read operation the block is validated by comparing the stored offset with the real offset, and the stored checksum with the real calculated checksum. The event execution time is the pure calculation part. If you run the test with multiple threads, it is the sum of the time of all threads.

2. Compare the results for the two operations (sequential, random). What are reasons for the differences?

   Based on the plots 3 and 4 it's clear that AWS wins random disk benchmark results, whereas GCP dominates in the sequential disk read benchmark.
   AWS instance provided almost the same experience in both random and sequential disk read performance evaluation. Nevertheless, the readings illustrate huge fluctuations in reading speed showing inconsistent quality of service.

   We use a standard persistent disk in our GCP instance.

   According to google cloud official documentation "Standard persistent

disks are efficient and economical for handling sequential read/write operations, but they aren't optimized to handle high rates of random input/output operations per second (IOPS). If your apps require high rates of random IOPS, use SSD persistent disks. SSD persistent disks are designed for single-digit millisecond latencies. Observed latency is application specific."[2].

For our AWS instance we use GP2 volume. According to AWS official website which is designed to offer single-digit millisecond latencies, deliver a consistent baseline performance of 3 IOPS/GB (minimum 100 IOPS) to a maximum of 16,000 IOPS, and provide up to 250 MB/s of throughput per volume[1].

4 General question

1. Compare the overall long-term measurement plots for the two platforms AWS and GCP. Name one type of application that you would expect to perform better on AWS, and one that would perform better on GCP, respectively. Shortly explain your decisions.

   In conclusion and based on the results we can safely say, AWS has substantial performance in the memory territory and great CPU performance as well when compared with GCP platform. Nevertheless, the collected performance readings are prone to change depending on the instance's hardware specifications and upgrades.
   AWS is more suitable for CPU and memory-hungry applications such as graphical applications.
   GCP offers more stable service and way better disk performance in terms of sequential read, which makes it more favourable for applications that read and write loads of data to the disk.

References

[1] Amazon ebs features - amazon web services. https://aws.amazon.com/ebs/features/: :text=General (Accessed on 11/20/2020).

[2] Storage options — compute engine documentation — google cloud. https://cloud.google.com/compute/docs/disks. (Accessed on 11/20/2020).