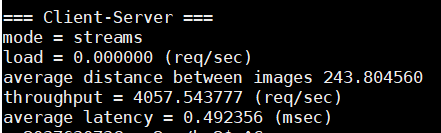
**Accelerators and Accelerated Systems**

**HW 2 – report**

1. **Streams Implementation**
   1. Throughput for **load = 0**: *4057.543777* req/sec



* 1. Number of run with various loads:

|  |  |  |
| --- | --- | --- |
| **throughput (reqs/sec)** | **latency (msec)** | **load** |
| 398.660753 | 2.673019 | 405.7544 |
| 1206.531737 | 0.964623 | 1262.347 |
| 1715.66365 | 0.707637 | 2118.94 |
| 2218.175088 | 0.563871 | 2975.532 |
| 2540.675385 | 0.502209 | 3832.125 |
| 2693.659502 | 0.478339 | 4688.717 |
| 2842.119098 | 0.457451 | 5545.31 |
| 2991.635421 | 0.440879 | 6401.902 |
| 3138.569202 | 0.422929 | 7258.495 |
| 3135.346558 | 0.426348 | 8115.088 |

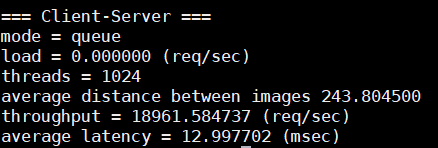
Observations:

1. as we increase the load we get better latency and throughput.
2. For lower loads, increase in load by improves latency exponentially. As we reach higher loads we can see that the latency improves linearly.
3. In every increase in load, throughput improves by smaller amounts (exponentially).

The results we are getting make sense. If we set the load to a low value, the gpu does not receive tasks fast enough to be efficient, i.e. it sometimes stays idle waiting for tasks to be sent to it or for finished calculation results to be collected from it. As we increase the load, tasks are sent to it fast which then improves the GPU's utilization, resulting in increase in throughput and decrease in latency. As we are sending more and more tasks in smaller intervals we are beginning to reach the GPU maximum compute capability, causing the performance to improve by smaller and smaller amounts.

1. **Producer-Consumer Implementation**

2.4.1. Throughput for **load = 0, #threads = 1024**: *18961.584737* reqs/sec.



2.4.2.

|  |  |  |
| --- | --- | --- |
| **throughput (reqs/sec)** | **latency (msec)** | **load** |
| 1646.689302 | 13.890456 | 1896.158 |
| 5620.734362 | 4.031477 | 5899.16 |
| 8415.277509 | 2.692663 | 9902.161 |
| 12007.39557 | 1.727553 | 13905.16 |
| 15644.21032 | 1.494067 | 17908.16 |
| 15237.97619 | 1.548804 | 21911.16 |
| 14374.48697 | 1.552689 | 25914.17 |
| 14796.64012 | 1.570418 | 29917.17 |
| 14093.87498 | 1.682353 | 33920.17 |
| 14027.64226 | 1.705405 | 37923.17 |