Structured Abstract

Name: Michael Luke Battle

Student Number: 200725640

Context:

Terapixel images offer a means to intuitively explore and present big data, however rendering such images is too computationally costly for conventional machines. Cloud supercomputing offers a route to render terapixel images efficiently, but as an emerging technology, should be rigorously evaluated prior to wider implementation.

Objective:

The aim of this project was to thoroughly analyse the terapixel image rendering process using cloud supercomputing, with 1024 GPU cores provided by Microsoft Azure. The primary goals of it are to fully comprehend the process and identify any inefficiencies.

Method:

Three data sets, produced as biproducts of the rendering process, were reformatted, combined and analysed, using R and the CRISP-DM analysis framework.

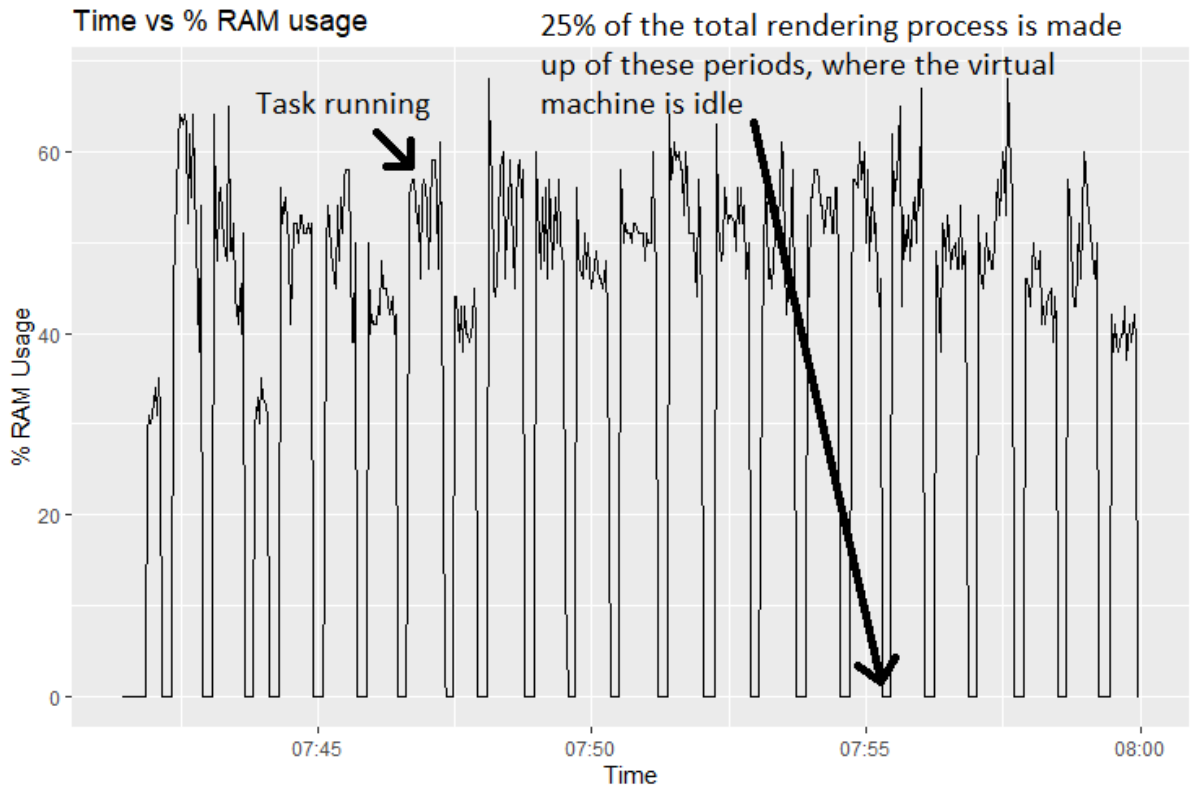
Results:

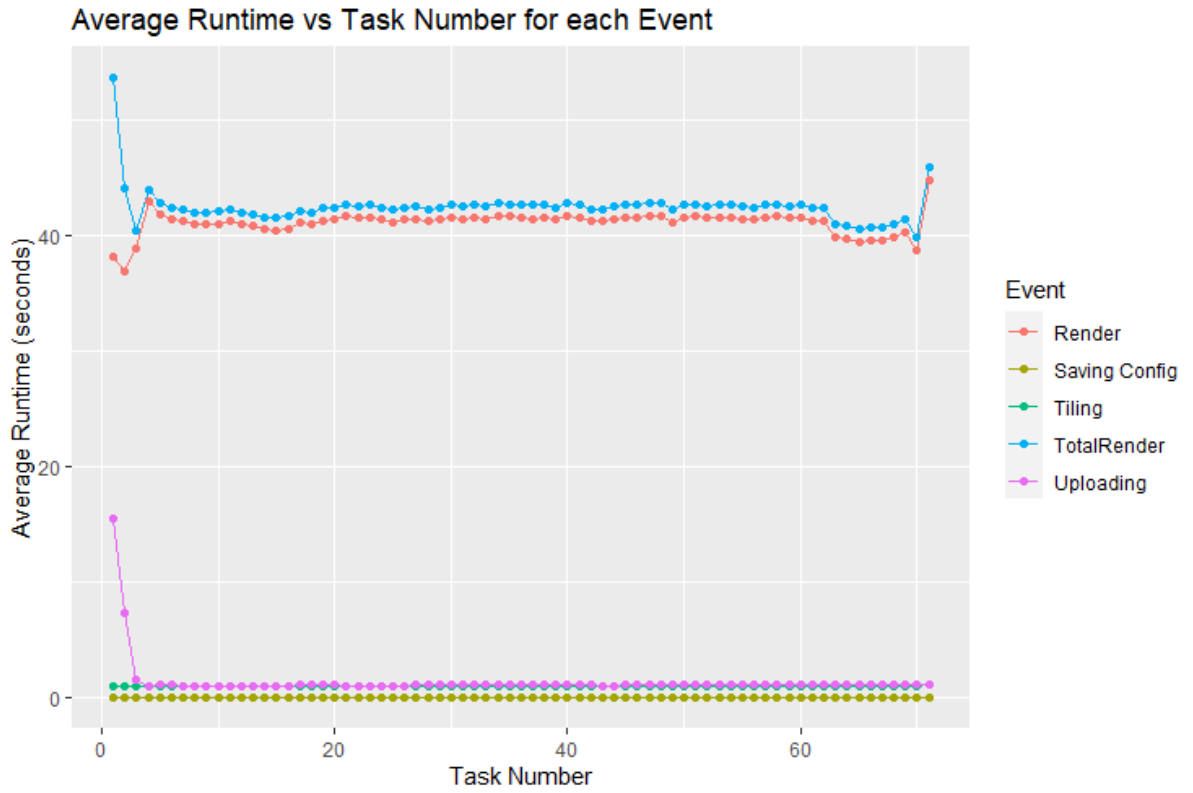
On around 80% of virtual machines, the first two tasks take approximately ten times longer to run the "Uploading" event than in the remainder of the process, and on average the GPU cores spend 25% of the total rendering time awaiting task scheduling. In terms of comprehension, positive correlations were identified between the render time per tile and both GPU temperature and power draw.

Novelty:

Previous work in this area has concerned the creation of the scalable cloud architecture to facilitate the rendering process. This work is built upon within this project, as the resulting rendering process is critically analysed in an endeavour to further our understanding of the operation, as well as improve it.

Key Images


Time vs % RAM usage
25% of the total rendering process is made up of these periods, where the virtual machine is idle
Task running
% RAM Usage
0
20
40
60
07:45
07:50
07:55
08:00
Time


Average Runtime vs Task Number for each Event
Average Runtime (seconds)
0
20
40
0
20
40
60
Task Number
Event
Render
Saving Config
Tiling
TotalRender
Uploading