Structured Abstract with Key Images

Context

With the use of cloud technologies on the rise, there is a need to find more information to improve efficiency, cut costs and power, and make the cloud even more useful. This can be done by analyzing the data from performance evaluations.

Objective

The objective of this study is to extract pertinent information from a dataset acquired from the application checkpoint and system metric output of a terapixel image generation. The aim is to identify the region that requires optimization focus.

Method

The data utilized for terapixel image processing was provided by Newcastle University, utilizing 1024 GPU nodes. The data was processed through exploratory analysis utilizing Python programming language within a Python Notebook, and version control was implemented through the use of GitHub.

Result

The research discovered that the optimization of GPU allocation, whereby high-end GPUs are designated for rendering events and mid to lower-tier GPUs are allocated for other tasks, can result in enhanced performance and cost reduction. Moreover, regions characterized by intricate configurations, such as edifices and thoroughfares, demand greater computational resources for visualization and warrant careful resource allocation.

Novelty

The existing research in this domain has predominantly centred on architectural remedies and the mitigation of power usage. The current research presents an innovative strategy that centres on cost reduction while simultaneously ensuring optimal performance through the execution of efficient resource management techniques.

Key Images

Power Consumption of GPUs

Box Plot of GPU Statistics