**Structured Abstract with Key Images**

**Context**

With the increasing use of cloud technologies, there is a need to find more information to enhance efficiency and further reduce cost and power. This can be done through performance evaluation using data analysis.

**Objective**

This project aims to extract information from the dataset created from the application checkpoint and system metric output from a production of a terapixel image to pinpoint the area on which optimization should be focused.

**Method**

The data was provided by Newcastle University and was created during the use of 1024 GPU nodes for terapixel image processing. It was wrangled, with exploratory data analysis done using Python in jupyter notebook along with the use of GitHub for version control.

**Result**