RESEARCH PROPOSAL

SUBMITTED TO THE

UNIVERSITY OF READING

Department of Computer Science

PHD Thesis

Author:
Kyle SPINDLER

Supervisor:[1]
Dr. Julian KUNKEL

October 25, 2018

Keywords: Microservice, Serverless, HPC

High-Performance Storage

[1]This thesis has been discussed with the listed supervisor.

Please also check `http://www.reading.ac.uk/computer-science/dcs-PhD-programmes.aspx`.

1 Motivation

Software Architecture invovles considering multiple characteristics such as separation of concerns, quality attributes (maintainability, scalability, loose coupling, high cohesion etc...) and architectural styles. Some architecture styles are more suited for performance while others are better at maintainability and loose coupling like microservices. Microservices is a very popular architecture that is used in many domains because of the benefits it offers.

Scientic codes suffer from good software engineering practices. HPC and store applications are typically tightly coupled to utilize the available resources efficiently. While it is claimed that this provides the best performance, the benefit and drawbacks of alternative software architectures for HPC software is not thoroughly investigated. Microservices, for example, provide a scalable architecture and ease the software development process by providing separation of concerns by applying techniques from Domain Driven Design. When deciding a software architecture not only performance and scalability matters, but also flexibility and maintainability.

In this regard, the HPC community struggles to recruit sufficient developesr to keep up with the development of software which can often be seen in important utility tools. For example, existing tools for pre/post-processing of HPC workflows and the analysis of HPC data are typically not the main focus of scientists and developers; hence, they are implemented in a way that shows limited scalability, i.e. are executed sequentially in bash scripts.

2 Research question

The main research question(s) that you want to address.
Understand the impact of modern day software architectures (microservices, event driven) has on HPC and particularly the climate/weather domain
The goal of this thesis is to see if HPC applications and storage systems can be redeveloped using modern day software architecture such as microservices with minimal or no overhead while gaining the benefits from the loosely coupled architecture.

1. What parts of the HPC application could benefit from microservices or other software architectures?
2. What

Example: The goal of this thesis is to understand and optimize the performance behavior for large-scale data accesses in the domain of climate and weather.
Now split the research goal into questions

This covers the research questions:

1. What workflows are limited by I/O?
2. Which I/O operations are typically performed?
3. Which optimizations are beneficial for the workflows on HPC systems?

3 Related work

How your thinking builds on any previous work.

Relevant work can be classified into: a) LaTeX studies, b) performance analysis in HPC, ...

LaTeX studies. It has been shown that blabla (Lamport, 1994).

Performance analysis.

4 Research methodology

What research methodology or techniques you may need to use

5 Required infrastructure

What facilities you are likely to require to conduct your research.

This research requires a supercomputer with more than 100 nodes to run experiments on.

6 Workplan

How the research can be completed in the time available. Provide a rough sketch over the runtime of your PhD
The following sketches the workplan for the different years of the PhD.

First year: setup of work environment, researching related work, writing the chapters introduction and related work of the thesis.

Second year:

Third year:

A Appendix

Add here any appendix, if needed

References

Lamport, Leslie (1994). *LATEX: a document preparation system: user's guide and reference manual*. Addison-wesley.