SET10108 Concurrent and Parallel Systems Coursework Specification 2018/19

Part 1: CPU-based Concurrency and Parallelism

Part 1 of the coursework is worth 40% of the total coursework mark. The submission date is **Sunday 28th of October (Week 7) online via Moodle by midnight**.

Part 1 of the coursework focuses on single-CPU concurrency and parallelism. You can use **multi-threading, OpenMP, algorithmic skeletons, and CPU-level parallelism.**

The aim of the coursework is to evaluate algorithm performance and attempt to improve the performance using parallel techniques. You must be systematic and document your attempts at parallelisation, investigating methods that may improve performance. **You will be graded on the rigour of approach, not how fast your application is**.

Algorithm

The algorithm provided is a block chain simulator. Moodle has the code to work on. You can tune the sequential algorithm as you wish as long as the behaviour is the same. Your aim is to improve performance via parallel techniques, and it is only in this area you will be assessed. You must investigate different configurations and measure performance using the techniques from the module (e.g. speedup and efficiency). Areas to explore include:

- Changing the difficulty of the required hash.
- Changing the data being stored.

The simulator has been tweaked to allow easier parallel performance. These changes have been highlighted. Otherwise the commenting in the code is minimal. The program uses a SHA-256 algorithm which can also be enhanced, although you cannot change the hash codes output from a given input. The hashing must operate correctly.

The application will produce the same hash code for a block on repeat runs. This is how you can check if your changes have broken the program.

Submission Requirements

The report for Part 1 must be 8 pages maximum (including appendices but not references). **Any page above greater than 8 will not be assessed**. The report must be in the following format:

- Font size 12pt.
- 1.5 line spacing.
- Margins 1 inch.

- Numbered pages.

The report must only contain the following sections:

- Title.
- Abstract.
- Introduction and Background.
- Initial Analysis.
- Methodology.
- Results and Discussion.
- Conclusion.
- References.

You must number the sections from Introduction and Background to Conclusion. Your submission must include your report as a PDF (**not a Word document**), and your code in a single ZIP file. **Do not submit executable files or build configuration files (e.g. Visual Studio project files)**. Penalties are:

Report not submitted as PDF 50% reduction of report marks.

Non-code files submitted 50% reduction of code marks.

All coursework must be demonstrated to the module leader in the practical of week 8. If the coursework is not demonstrated it will receive a grade of 0.

Marking Scheme for Part 1

An overview of the marking scheme is below:

Component	**Description**	**Mark**
Initial analysis	Initial analysis of the base-line performance of the application and likely places that can be parallelised.	5
Methodology used	Description and justification of the approach used and its overall suitability and rigour.	5
Results presented	Suitable performance analysis and testing documentation for the problem, including quality of presentation of the results.	10
Conclusions drawn	Level of discussion and appropriateness of the conclusions drawn based on the results gathered.	5
Report quality	Adherence to report requirements and overall quality of discussion and presentation.	10
Application code	The quality of the code generated at all stages of the coursework.	5

A breakdown of these marks is given in the following table:

Section	Feedback	Grade
Initial Analysis	Gathered baseline data of the application. Baseline results include manipulation of variables to determine their impact on execution time. This is used to analyse the algorithm. Measurements and tools are also used to determine likely bottlenecks down to the code level.	A+ (5 marks)
	There are baseline measurements and an examination of the variables. This has helped identify code to work on. No actual algorithmic analysis.	A (4 marks)
	Baseline measurements taken including experimentation on variables. Specific conclusions on the complexity or the code to work on not made.	B (3 marks)
	Some baseline measurements or use of a profiler to identify the code to work on.	C (2 marks)
	Some hand analysis of the code but not any measurements.	F (1 mark)
	Not done.	F (0 marks)
Methodology	The general approach is well defined, and the technological approach is likewise defined. Suitable justification based on initial analysis for why this approach was taken also made.	A+ (5 marks)
	The general approach is well defined, and the technology is discussed. No justification for the choices made.	A (4 marks)
	The general approach is defined, or the technological approach is defined and why it is being used.	B (3 marks)
	The technology is defined and likely how it will be used. No discussion on the general experimental approach.	C (2 marks)
	A limited discussion on the general experimental approach or technological approach provided.	F (1 mark)
	Not provided, or just a mention of the technology.	F (0 marks)
Results	Suitably comprehensive set of results gathered, including investigation of different variables. The results are presented well. Correct charts are used, and these are of a high-standard, including correct scaling. Speedup and efficiency measures provided for all experiments using the correct values.	A+ (10 marks)
	As A+, but likely not a fully comprehensive set of results, or some results presentation issues.	A (8-9 marks)
	Something is missing from the results. Either the charting is not done well, or speedup or efficiency measures are missing. It may also be that only limited results have been gathered.	B (6-7 marks)
	Speedup and efficiency measures missing or lacking, or the presentation and results are limited.	C (4-5 marks)
	Minimal results provided, and presentation is poor. The standard metrics of efficiency and speedup are not provided.	F (2-3 marks)
	Either results not provided or the wrong part of the application parallelised leading to invalid results.	F (0-1 marks)
Conclusions	Strong conclusions drawn, related to the results gathered and the work undertaken. Good summarisation of the work and results, highlighting key points. Future work and limitations mentioned.	A+ (5 marks)
	Conclusions drawn, and summarisation of both the work undertaken and the results generated. Possibly future work and limitations mentioned.	A (4 marks)
	Conclusions drawn, and some summarisation of the work undertaken or the results.	B (3 marks)
	Some conclusions drawn, but no summarisation of the work.	C (2 marks)
	Minimal conclusions made, or just a summarisation of the work.	F (1 mark)
	No conclusions made, or just a statement on the work.	F (0 marks)
Report Quality	Penalty for each requirement not met	1 mark
	Excellent overall presentation and discussion throughout	(10 marks)
	Excellent discussion but some presentation issues	(9 marks)
	Generally very good discussion throughout but some presentation issues	(8 marks)
	Discussion and presentation could use work in some places	(6-7 marks)
	Discussion and presentation poor and lacking thread of ideas through text	(5 marks)
Application Code	Different code files submitted based on different approaches taken. Code is of a high-quality throughout. Suitable comments are made highlighting where changes were made and describing what they do.	A+ (5 marks)
	Commenting is descriptive on the changes made. There may also be some comments highlighting areas of understanding.	A (4 marks)
	Commenting highlights the changes made to the original sequential version, but is not descriptive.	B (3 marks)
	Some general commenting undertaken of where the work is but limited. Code may be untidy in places.	C (2 marks)
	Code submitted has no comments and is just a simply parallelised version of the sequential source. Code is likely messy and inconsistent.	F (1 mark)
	No code submitted.	F (0 marks)

Part 2: General Concurrency and Parallelism

Part 2 of the coursework is worth 60% of the total coursework mark. The submission date is **Sunday 9th of December (week 13) online via Moodle by midnight**.

In Part 2 of the coursework you choose a project to implement and then parallelise, undertaking performance analysis as Part 1. The aim is to investigate the problem and technologies used, and evaluating their effectiveness in context. You can now choose any techniques you wish, but **you must demonstrate the application on a machine in D2 (the Code Lab)**. This is because the module must be able to run your application.

Projects

There are three projects you can choose to undertake, or you may define a custom one. The projects range in difficulty from easy to hard, and your marks will reflect the complexity of the project you undertake.

Easy ray tracer. A simple 100 line C ray tracer can be found at `http://www.kevinbeason.com/smallpt/`. You can use this as a basis for your baseline code and parallelise accordingly. **You must save the image from the ray tracing to prove that it works.**

Medium N-body simulation. Information about the N-body problem is available from Wikipedia. There are solutions available, but you will need a sequential one that you can parallelise. **You must produce a visual output that proves the N-body simulation works correctly.**

Hard JPEG compression. JPEG compression has a number of features that can be manipulated to examine the impact on performance. As with the N-body simulation, solutions exist online.

Custom discuss your idea with the module leader if you want to attempt something different.

You are not expected to use every parallel technique. Pick an appropriate technology and a strategy based on the nature of the problem. You may use CPU-based, GPU-based, or distributed techniques. You may mix these as well. You are not limited to C++ at this point and may use any language you are comfortable with if you can justify your choice to the module leader.

Submission Requirements

The report for Part 2 must be 12 pages maximum (including appendices but not references). **Any page above greater than 12 will not be assessed**. The report must be in the following format:

- Font size 12pt.
- 1.5 line spacing.
- Margins 1 inch.
- Numbered pages.

The report must only contain the following sections:

- Title.

- Abstract.
- Introduction and Background.
- Initial Analysis.
- Methodology – including justification of approach based on algorithm.
- Results and Discussion.
- Conclusion.
- References.

You must number the sections from Introduction and Background to Conclusion. Your submission must include your report as a PDF (**not a Word document**), and your code in a single ZIP file. **Do not submit executable files or build configuration files (e.g. Visual Studio project files)**. Penalties are:

Report not submitted as PDF 50% reduction of report marks.

Non-code files submitted 50% reduction of code marks.

An initial prototype of your coursework must be demonstrated to the module leader in the practical of week 13. If the coursework is not demonstrated it will receive a grade of 0.

Marking Scheme for Part 2

An overview of the marking scheme is below:

Component	**Description**	**Mark**
Initial analysis	Initial analysis of the base-line performance of the application and likely places that can be parallelised.	5
Methodology used	Description and justification of the approach used and its overall suitability and rigour.	10
Results presented	Suitable performance analysis and testing documentation for the problem, including quality of presentation of the results.	10
Conclusions drawn	Level of discussion and appropriateness of the conclusions drawn based on the results gathered.	5
Report quality	Adherence to report requirements and overall quality of discussion and presentation.	10
Application code	The quality of the code generated at all stages of the coursework.	10
Complexity	Awarded based on the difficulty of the problem undertaken and the overall challenge of the approach used.	10

A breakdown of these marks is given in the following table:

Section	Feedback	Grade
Initial Analysis	Gathered baseline data of the application. Baseline results include manipulation of variables to determine their impact on execution time. This is used to analyse the algorithm. Measurements and tools are also used to determine likely bottlenecks down to the code level.	A+ (5 marks)
	There are baseline measurements and an examination of the variables. This has helped identify code to work on. No actual algorithmic analysis.	A (4 marks)
	Baseline measurements taken including experimentation on variables. Specific conclusions on the complexity or the code to work on not made.	B (3 marks)
	Some baseline measurements or use of a profiler to identify the code to work on.	C (2 marks)
	Some hand analysis of the code but not any measurements.	F (1 mark)
	Not done.	F (0 marks)
Methodology	The general approach is well defined, and the technological approach is likewise defined. Suitable justification based on initial analysis for why this approach was taken also made.	A+ (10 marks)
	The general approach is well defined, and the technology is discussed. No justification for the choices made.	A (8-9 marks)
	The general approach is defined, or the technological approach is defined and why it is being used.	B (6-7 marks)
	The technology is defined and likely how it will be used. No discussion on the general experimental approach.	C (4-5 marks)
	A limited discussion on the general experimental approach or technological approach provided.	F (2-3 mark)
	Not provided, or just a mention of the technology.	F (0-1 marks)
Results	Suitably comprehensive set of results gathered, including investigation of different variables. The results are presented well. Correct charts are used, and these are of a high-standard, including correct scaling. Speedup and efficiency measures provided for all experiments using the correct values.	A+ (10 marks)
	As A+, but likely not a fully comprehensive set of results, or some results presentation issues.	A (8-9 marks)
	Something is missing from the results. Either the charting is not done well, or speedup or efficiency measures are missing. It may also be that only limited results have been gathered.	B (6-7 marks)
	Speedup and efficiency measures missing or lacking, or the presentation and results are limited.	C (4-5 marks)
	Minimal results provided, and presentation is poor. The standard metrics of efficiency and speedup are not provided.	F (2-3 marks)
	Either results not provided or the wrong part of the application parallelised leading to invalid results.	F (0-1 marks)
Conclusions	Strong conclusions drawn, related to the results gathered and the work undertaken. Good summarisation of the work and results, highlighting key points. Future work and limitations mentioned.	A+ (5 marks)
	Conclusions drawn, and summarisation of both the work undertaken and the results generated. Possibly future work and limitations mentioned.	A (4 marks)
	Conclusions drawn, and some summarisation of the work undertaken or the results.	B (3 marks)
	Some conclusions drawn, but no summarisation of the work.	C (2 marks)
	Minimal conclusions made, or just a summarisation of the work.	F (1 mark)
	No conclusions made, or just a statement on the work.	F (0 marks)
Report Quality	Penalty for each requirement not met	1 mark
	Excellent overall presentation and discussion throughout	(10 marks)
	Excellent discussion but some presentation issues	(9 marks)
	Generally very good discussion throughout but some presentation issues	(8 marks)
	Discussion and presentation could use work in some places	(6-7 marks)
	Discussion and presentation poor and lacking thread of ideas through text	(5 marks)
Application Code	Different code files submitted based on different approaches taken. Code is of a high-quality throughout. Suitable comments are made highlighting where changes were made and describing what they do.	A+ (10 marks)
	Commenting is descriptive on the changes made. There may also be some comments highlighting areas of understanding.	A (8-9 marks)
	Commenting highlights the changes made to the original sequential version, but is not descriptive.	B (6-7 marks)
	Some general commenting undertaken of where the work is but limited. Code may be untidy in places.	C (4-5 marks)
	Code submitted has no comments and is just a simply parallelised version of the sequential source. Code is likely messy and inconsistent.	F (2-3 mark)
	No code submitted.	F (0-1 marks)
Complexity	A challenging approach to parallelisation undertaken incorporating multiple technologies. Different algorithmic approaches taken to try and solve the problem, and these are also parallelised using combinations of technologies. A comparison of technologies also undertaken.	A+ (10 marks)
	A challenging approach and problem, and a combination of technologies likely explored but no in-depth comparison.	A (8-9 marks)
	A challenging problem and approach taken, and some comparison between technologies that have been used.	B (6-7 marks)
	A challenging problem or approach taken, but likely only using one technology.	C (4-5 marks)
	Simple problem undertaken, or a simple approach taken. Likely only one technology type examined.	F (2-3 marks)
	Simple approach taken to a simple problem. Likely only one technology type examined.	F (0-1 marks)