CAS 720: Foundations of Future Resilient Databases
- Course Outline

Jelle Hellings

Department of Computing and Software
McMaster University

1 Administrative Details

Course title Foundations of Future Resilient Databases

Course number CAS 720

Course categorization Systems.

Term

Lecturer Jelle Hellings (✉ jhellings@mcmaster.ca, ☁ https://jhellings.nl)

Department of Computing and Software, McMaster University.

Please use e-mail for *all* communication with the lecturer.

Lectures Two lectures per week (32 lectures in total

- ▶
- ▶

Website Avenue to Learn (☁ https://avenue.mcmaster.ca) for announcements, course material, assignments, assignment submissions, this course outline, and miscellaneous materials.

Office hours Regular office hours will be held Tuesdays and Wednesdays (1:00PM–3:00PM). Meetings outside the regular office hours are by e-mail appointments and can be either face-to-face or virtual (via Microsoft Teams).

2 Course Description

The field of databases is ever evolving. Recent directions include the development of resilient database systems based on blockchain technology that can deal with failures and enable cooperative data management. In this course, we will provide an introduction to this area, starting with a basic treatment of consensus-based transaction processing and with a further focus on selected topics.

In specific, this course will cover selected topics on distributed database systems with a main focus on resilient database systems. The topics will include: (1) transaction processing in distributed databases, (2) consensus protocols (PBFT, Proof-of-Work), (3) Merkle trees and other ledger-supporting data structures, and (4) cluster-sending and other primitives for cooperation between resilient systems.

3 Course Structure

The course consists of three parts: introduction lectures, paper study, and the course project.

3.1 Introduction Lectures

As the name suggests, the introduction lectures will introduce the students to database systems, distributed systems, fault-tolerant (resilient) systems, and the fundamental techniques powering them. These introductory lectures will sufficiently introduce the background necessary to participate in this course, to independently study the selected papers, and to start on the course project.

3.2 Paper Study

During paper study, we will take a look at selected topics from research papers. This format will represent the majority of all course content. We refer to Section ?? for an overview of the research papers we will look at. The selection of research papers will cover core results, work of general interest, and recent developments.

Short summary. In preparation to a paper study session, *all students* are required to read the studied paper and prepare and **submit a short summary** of the paper to Avenue to Learn. The deadline for submission of a short summary of a paper is the Sunday before the paper study for that paper. E.g., on Sunday October 2, the summaries for three papers ([**blockdist**, **aca_ped**, **3pc**]) are due. Each short summary should follow the template of Appendix ??. Write the summary ***independently and using your own words***, avoid general comments (e.g., weak motivation, poor writing, and so on), and be as specific as possible.

Presentation. In addition, each student will **present** one of the paper study papers during the semester (the exact paper load can change, depending on the number of students). Students must also share their presentation slides via Avenue to Learn.

The presentation should provide a good introduction to the content of the paper and serve as the basis for a discussion on the paper. Presentations should be around 35–45 minutes and should cover the main technical details of the paper. Furthermore, presentations should address the *why* of the paper (motivation) and the *what* of the paper (outcome):[1]

Why was the paper written? Motivation	Context	Why is the topic important?
	Need	What topic is studied in the paper?
	Task	What was done to address the topic?
	Object	What does the paper do (technical details)?
What are the results? Outcome	Findings	What does the paper reveal?
	Conclusions	What do the findings mean?
	Perspectives	What are the steps beyond the paper?

Presentation can be based on external material (e.g., if presentation slides are provided online by the authors). In that case, students need to ***unambiguously cite their sources*** and assure that they understand what they are presenting. Students are encouraged to discuss their paper and draft of their presentation with the instructor via an Office Hour meeting the week before their presentation (please make an appointment to do so to allow the instructor to revisit the paper beforehand).

Students must coordinate with the instructor for the paper(s) they wish to present out of the available papers (see Section ??). We use a first-come, first-serve policy when assigning papers. All paper preferences must be sent to the instructor before September 27. If no preferences are communicated by the due date, a random assignment will be applied.

---

[1]The table providing a high-level structure of a presentation is based on *https://www.principiae.be/book/* by Jean-luc Doumont.

3.3   Course Project

A big part of this course is the *course project* on a topic broadly related to resilient databases (e.g., consensus protocols, blockchains, or distributed systems). All projects are done by groups of 3-5 students (the exact number will depend on the number of students). One can choose out of three types of projects:

▶ A *research project* in which the group designs and evaluates a novel (part of) a resilient database.

The core objective of a research project is the *study of a novel problem* related to resilient databases. Your project must not only motivate and analyze the problem, but also propose a solution to the problem. A research project requires an *evaluation* of the solution in order to demonstrate *feasibility*. The evaluation of your solution can include performance modeling, performance measurements on an implementation (e.g., in case of an algorithm), and comparisons with pre-existing solutions to the studied problem.

Examples of topics for research projects are the design of a novel approach towards data replication, consensus, or sharding in resilient databases.

▶ A *whitepaper project* in which the group studies a specific use-case of resilient databases.

The core objective of a whitepaper project is the *application of resilient database technology* to solve a practical problem. Your project must identify a *use-case*, analyze the use-case, identify practical challenges with respect to that use-case, and design a system that can tackle these challenges. The final result of a whitepaper project is a *proof-of-concept* prototype that demonstrates the *feasibility* of your design. As part of the project, you will need to evaluate the design and prototype by showing how they tackle the identified challenges.

Examples of topics for whitepaper projects are the usage of *blockchain technology* to provide data integration between multiple stakeholders in an industry, to provide data traceability in an industry for regulatory purposes, and so on.

▶ An *evaluation project* in which existing works related to resilient databases are evaluated.

The core objective of an evaluation project is the evaluation of existing works. Your project will choose existing works (software artifacts, source code, or a paper), analyze the works, identify what problems are solved by the works, understand how the works solve these problems, and evaluate the quality of the solution (how good does the solution solve the original problem). An evaluation project should have an implementation component (e.g., implemented algorithms when evaluating papers, implementation of a benchmark environment when evaluating software artifacts or source code), should clearly identify strengths and weaknesses of the evaluated works, and should identify possibilities for future research steps based on the evaluated works.

Evaluation projects can focus on existing blockchain platforms and fabrics (e.g., Hyperledger Fabric, ResilientDB, Bft-SMaRt, etc); on existing consensus protocols (e.g., implementing and comparing existing protocols); or on underlying techniques (e.g., a study of practical threshold signatures or information dispersal algorithms).

Within the project, the role and contributions of each group member must be clearly articulated throughout the project. The project is partly evaluated via *self-evaluation* and *peer-evaluation* (see Appendix **??** for the evaluation form used during the mid-term and final evaluation).

All projects will have an initial evaluation of the project proposal (due October 10), a mid-term evaluation (presentation and mid-term report, due November 15), and a final evaluation (presentation and final report, due December 20). Project improvements made based on feedback (during progress meetings, on the project proposal, or during the mid-term evaluation) will be reflected in the overall grade of the project. The project proposal, mid-term, and final reports should be prepared in LaTeX and follow the format of a *formal paper* (e.g., such as those studied during the paper study component of the course). We provide a paper template on Avenue to Learn that can be used as the starting point for these LaTeX documents.

On group work. Effective collaborating with others is a crucial skill to develop for any future career (both academics and industry). To promote effective collaboration, we encourage every group member to *act responsibly*, *actively work* on providing a healthy work environment, and put a *fair effort* into their contributions to the project.

We believe it to be important to learn how to resolve our differences and we ask all group members to be kind to each other, to listen to each other, to be open to what others have to say, and to be supportive of each other. Once a group is formed, we expect the group to work in harmony and try to resolve any internal conflicts in a healthy and responsible manner. In normal circumstances, it should not be necessary for the instructor to mediate internal group affairs. *Do reach out* to the instructor if you feel things are going wrong.

4 Grading

The grade is determined by the *paper study* (30%) and by the *course project* (70%). The detailed grade breakdown is as follows:

Paper study (35%).

- ▶ 15% for the paper summaries (due on Sundays); and
- ▶ 20% for the paper presentation.

Course project (65%).

- ▶ 10% for the project proposal (due October 10);
- ▶ 15% for the mid-term evaluation (November 15);
- ▶ 10% for the final presentation (December 20);
- ▶ 30% for the final report (due December 20).

Bonus grades. Students that find errors or have strong suggestions for improvements in either the book [**book**] or in the slides of the *Introductions* lectures are encouraged to share their findings. Minor bonus grades will be awarded when the findings of the student lead to the improvement of either the book or the slides.

Late submission policy. All due dates are *strict* at midnight and any work submitted after a grace period of five hours after midnight will *not be graded*. Contact the instructor before the deadline if you have technical difficulties while submitting your work. Deadline extensions can be granted by the instructor in case of exceptional circumstances.

5 Plagiarism and Academic Dishonesty

To facilitate a healthy learning environment for all, we require integrity of the students. Any suspected breaches of integrity will be handled according to the Academic Integrity Policy.

Do not submit or present work of *others* as your own. All parts of the work you submit or present for this course *must* be your own work, in your own words, and be based on your own ideas and conclusions. If parts of your work are based on the works of others, then these parts must be accompanied by *unambiguously citations* of their sources. Do not wholeheartedly copy or paraphrase large sections of other works, unless clearly demarcated as such. In case of doubt, contact the instructor *before handing in your work*.

<div align="center">

Plagiarism is a *serious academic offense* and will be handled accordingly.
All suspected cases of academic dishonesty will be reported to the *Office of Academic Integrity*.

</div>

Paper summaries and paper presentations are *individual assignments* and should be made as such. By submitting or presenting your work, you are certifying that the submitted or presented work is made *by you* in accordance with the above *plagiarism and academic dishonesty* statement. All your works can be subject to automated and manual plagiarism checking.

6 Course Schedule and Course Material

The schedule of this course is subject to change (e.g., based on the number of students that are taking the course and the number of project groups). Next, one can find the preliminary course schedule:

Date	Topic	Required	Course Material Additional
Sep. 6	Course Organization. Introductions: Distributed Systems.	[**book**]	[**distbook**, **distdb**, **capproof**, **cap12**]
Sep. 7	Introductions: Resilient Systems.	[**book**]	[**distalgo**, **asyncbc**, **flp**, **generals**]
Sep. 13	Introductions: PBFT (part 1)	[**book**]	[**pbftj**, **jpoe**]
Sep. 14	Introductions: PBFT (part 2)	[**book**]	[**pbftj**, **jpoe**]
Sep. 20	Introductions: Performance Modeling for Consensus	[**book**]	
Sep. 21	Introductions: Cooperating Resilient Systems	[**csp**], [**book**]	
Sep. 27	Introductions: Sharding	[**byshard**]	
Sep. 28	Introductions: Proof-of-Work, Merkle trees, Bitcoin, and Ethereum	[**bitcoin**]	[**ethereum**]
Oct. 4	Paper study: "Blockchains from a distributed computing perspective" and "Bitcoin's Academic Pedigree"	[**blockdist**] [**aca_ped**]	
Oct. 5	Paper study: "Nonblocking Commit Protocols"	[**3pc**]	
Oct. 11	Paper study: "Conflict-Free Replicated Data Types"	[**crdt**]	
Oct. 12	Paper study: "Efficient Byzantine Fault-Tolerance"	[**minbft**]	
Oct. 18	Paper study: "Scalable Secure Storage when Half the System Is Faulty"	[**scale_ida**]	
Oct. 19	Paper study: "Cross-Chain Deals and Adversarial Commerce"	[**deal**]	
Oct. 25	Paper study: "HotStuff: BFT Consensus with Linearity and Responsiveness"	[**hotstuff**]	
Oct. 26	Project: Progress Meeting		
Nov. 1	Paper study: "Making Byzantine Consensus Live"	[**livebft**]	
Nov. 2	Paper study: "Small Byzantine Quorum Systems"	[**smallbqs**]	
Nov. 8	Paper study: "Algorand: Scaling Byzantine Agreements for Cryptocurrencies"	[**algorand**]	
Nov. 9	Project: Progress Meeting		
Nov. 15	Project: Mid-Term Presentations		
Nov. 16	Project: Mid-Term Progress Meeting		
Nov. 22	Paper study: "In Search of an Understandable Consensus Algorithm"	[**raft**]	
Nov. 23	Paper study: "Zab: High-Performance Broadcast for Primary-Backup Systems"	[**zab**]	
Nov. 29	Paper study: "Narwhal and Tusk: A DAG-Based Mempool and Efficient BFT Consensus"	[**narwhal**]	
Nov. 30	Project: Progress Meeting		
Dec. 6	Paper study: "The Byzantine Generals Strike Again"	[**byzgena**]	
Dec. 7	Project: Progress Meeting		
Dec. 13	Paper study: "The Honey Badger of BFT Protocols"	[**honeybadger**]	
Dec. 14	Project: Progress Meeting		
Dec. 20	Final Project Presentations		
Dec. 21	Project: Final Progress Meeting		

All of the course material referred to in the above table is available free of charge either online or via the

McMaster University Library. Contact the instructor if you have troubles getting a hold on any of the course material. Next, we provide the details on the referred course material:

A   Paper Study: Paper Summary Template

Paper quality: Choose one of: *excellent, good, fair, poor, bad.*

Description: by your own opinion, what is the *general quality* of the paper. E.g., does the paper present a clear idea, is the complexity of the paper proportional to the complexity of the subject matter, is the paper well-written the paper, and so on.

Personal interest: Choose one of: *extremely, very, moderately, slightly, not at all* interested.

Description: how interesting is the paper *for you.* Note that an excellent paper can be on a topic that is of no interest for you, while a bad paper can inspire many new ideas (making it very interesting to *you*)!

Summary

Description: provide a 100–200 word *summary* of the paper *in your own words.*

Strong points

Description: list three major strong points of the paper *in your own words.*

1. _____.
2. _____.
3. _____.

Opportunities for improvement

Description: list three major opportunities for improvement for the paper *in your own words.*

1. _____.
2. _____.
3. _____.

B Project Self-Evaluation and Peer-Evaluation

Student name: _____.

Description: *your* name.

Project topic: _____.

Description: the topic of *your* project.

Project Evaluation

Project grade: _____.

Description: what is the letter grade *you* believe best represents the current quality of the work done for and results of the project.

Strong points of the project

Description: list the three things that went really well during the project. This covers both technical aspects (presentations, writing, implementation, analysis, and so on) and the way in which the group collaborated and functioned as a whole.

   1. _____.

   2. _____.

   3. _____.

Opportunities for improvement for the project

Description: list three ways in which the project can be improved.

   1. _____.

   2. _____.

   3. _____.

Self-Evaluation and Peer-Evaluation

Description: fill this out for *each student* in your project group, *including yourself*.

The goal of *self-evaluation* and *peer-evaluation* is to support effective collaboration, to support learning, and to support each other's growth and development. As such we encourage everyone to provide <u>constructive feedback</u> and try to be objective, fair, and honest in your evaluation of both your own contributions and the contributions by others.

We will *never* directly share your evaluation with other students. We will share high-level *anonymized feedback* with other students, however.

Evaluation of: _____.

Description: name of the student you are evaluating.

Contribution Quality: Choose one of: *excellent, above average, average, below average, bad.*

Description: what do you believe is the *general quality* of the project contributions by the evaluated student?

Strong points

Description: list three things that the evaluated student did really well during the project.

1. _____.
2. _____.
3. _____.

Opportunities for improvement for the project

Description: list three ways in which the evaluated student can improve their contribution to the project.

1. _____.
2. _____.
3. _____.