COMP90056 Stream Computing and Applications

Credit Points:	12.5
Level:	9 (Graduate/Postgraduate)
Dates & Locations:	This subject is not offered in 2015.
Time Commitment:	Contact Hours: 36 hours (1 two-hour lecture per week and 1 one-hour tutorial/lab per week) Total Time Commitment: 200 hours
Prerequisites:	**Subject** / **Study Period Commencement:** / **Credit Points:** COMP90015 Distributed Systems — Semester 1, Semester 2 — 12.50
Corequisites:	None
Recommended Background Knowledge:	Java programming language, design of algorithms, distributed systems
Non Allowed Subjects:	None
Core Participation Requirements:	For the purposes of considering request for Reasonable Adjustments under the Disability Standards for Education (Cwth 2005), and Student Support and Engagement Policy, academic requirements for this subject are articulated in the Subject Overview, Learning Outcomes, Assessment and Generic Skills sections of this entry. It is University policy to take all reasonable steps to minimise the impact of disability upon academic study, and reasonable adjustments will be made to enhance a student's participation in the University's programs. Students who feel their disability may impact on meeting the requirements of this subject are encouraged to discuss this matter with a Faculty Student Adviser and Student Equity and Disability Support: http://services.unimelb.edu.au/disability
Contact:	email: **lkulik@unimelb.edu.au (mailto:saurabhg@unimelb.edu.au)**
Subject Overview:	**AIMS** With exponential growth in data generated from sensor data streams, search engines, spam filters, medical services, online analysis of financial data streams, and so forth, there is demand for fast monitoring and storage of huge amounts of data in real-time. Traditional technologies were not aimed to such fast streams of data. Usually they required data to be stored and indexed before it could be processed. Stream computing was created to tackle those problems that require processing and classification of continuous, high volume of data streams. It is highly used on applications such as Twitter, Facebook, High Frequency Trading and so forth. The Stream computing course will interest students who want to learn more about real-time processing and its applications. It will be taught both from a theoretical and practical point of view. The course will cover underlying fundamentals of stream processing systems, particularly architectural issues and algorithms for stream processing, mining and analysis. It will also include tutorials on how to develop and deploy applications into platforms such as IBM InfoSphere Streams®. **INDICATIVE CONTENT** # Why stream processing is important # Data streams model # Data streams algorithms: Sampling, sketching, distinct items, frequent items, etc. # Data streams synopses: Histograms, sketches, wavelets, etc. # Stream processing platforms: Infosphere Streams, Storm, Spark Streaming, etc. # Data streams mining: Classification, clustering, etc.

Learning Outcomes:	**INTENDED LEARNING OUTCOMES (ILO)** On completion of this subject the student is expected to: 1 Differentiate between stream computing models and conventional computing models 2 Understand stream computing algorithms and how to apply them in real world problems 3 Customise and create stream algorithms based on application requirements 4 Design, develop and deploy stream computing based applications
Assessment:	Two programming assignments, requiring approximately 25 - 30 hours of work each (20%) due in week 3 and 6 Four class quizzes, part of the tutorial/lab classes (10%), due in week 7, 9, 10 and 11 Final project with 20 min presentation, requiring approximately 50-55 hours of work (40%), due in week 12 Two hour end of semester exam (30%). Intended Learning Outcomes (ILOs) 1 and 2 are addressed in all components of assessment. ILO 3 and 4 are assessed in the end of semester exam and the final project.
Prescribed Texts:	None
Breadth Options:	This subject is not available as a breadth subject.
Fees Information:	Subject EFTSL, Level, Discipline & Census Date, http://enrolment.unimelb.edu.au/fees
Generic Skills:	On completion of this subject students should have the following skills: # Ability to undertake problem identification, formulation and solution # Capacity for independent critical thought, rational inquiry and self-directed learning # Profound respect for truth and intellectual integrity, and for the ethics of scholarship.
Notes:	**LEARNING AND TEACHING METHODS** The subject involves 1 two-hour lecture per week followed by a 1-hour workshop. Weekly, workshop problems are assigned and discussed during workshop hour. As the subject relies heavily on learning by practice, we have a good load of programming exercises as part of workshops and assignments. Students will work individually or on groups of two to implement algorithms and problems described during lectures and in the workshop. **INDICATIVE KEY LEARNING RESOURCES** The subject uses online reading materials (provided as recommend readings weekly) and online discussion forum. It offers access to slides, book chapters and relevant papers. **CAREERS /INDUSTRY LINKS** Stream processing is becoming more important as the world goes instrumented. Collecting and analysing data became easier and cheaper. One can access the importance of stream processing by looking the number of stream processing platforms being created recently. Stream processing is a key part of the Massive Data Analytics trend.
Related Course(s):	Master of Information Technology Master of Information Technology
Related Majors/Minors/ Specialisations:	MIT Computing Specialisation