

## Chapter 12

# Automating a Massive Online Course With Cluster Computing

**Timothy C. Haas**

*University of Wisconsin–Milwaukee, USA*

### ABSTRACT

*Before massive numbers of students can take online courses for college credit, the challenges of providing tutoring support, answers to student-posed questions, and the control of cheating will need to be addressed. These challenges are taken up here by developing an online course delivery system that runs in a cluster computing environment and is designed to support the delivery of a course having 10K or more students. This delivery system enhances synchronous and asynchronous lectures, provides an online intelligent tutoring system, and detects plagiarism. The free software system is shown to provide fast response times when run on a mid-range cluster computer. The system's automatic plagiarism detection system is shown to be able to detect multiple authors of course assignments when used to analyze the work of actual online students. Use of this system on a large scale would allow most colleges to reduce their faculty size by 20 to 40%.*

### INTRODUCTION

I describe a way to offer an online course to a massive number of matriculated students. Such students would pay a tuition fee for the course and receive a grade along with course credit on their transcript – in other words, the envisioned course would have exactly the same standing as any course taken on-campus for credit towards a degree. In this article, such a course is called a *for-credit massive online course* (MOC), i.e., a for-credit alternative to a *massive open online course* (MOOC) (The Chronicle, 2012).

MOCs that were inexpensive, say no more than \$100.00 in course tuition but included online tutoring and real-time response to student-posed questions would be attractive to students thinking of matriculating but who are either unable to afford the tuition of most present-day college courses or unable to attend on-campus courses. There may be large numbers of such students.

There are however, several challenges to be overcome before such MOCs can be offered. These include the need for tutoring support, real-time responses to student-posed questions, and the mitigation

DOI: 10.4018/978-1-5225-8057-7.ch012

of student plagiarizing of exams and other course assignments. It is argued herein that many of these challenges are in-part due to the formidable scale problem that a MOC presents and hence could be addressed by running the software that delivers a MOC within a cluster computing environment.

## **Business Models**

The business model for offering massive online courses that do not count towards a degree is difficult to justify for all but the most elite universities. As is well-known, revenue projections for MOOCs are qualitatively different than current for-credit courses offered either on-campus or online. Such ambitious projections are proving to be over-stated (Gütl et al. (2014)).

It is argued here that revenue projections are less speculative for MOCs. For example, if 10K students take a MOC during a single semester and each pays \$100.00 to do so, revenue from this course is \$1 million. If no more than a handful of staff were needed to maintain the course, profit margins could become very large. Because the MOC would consist of only matriculated students, dropout rates may be lower than the current 60 to 80% of certificate-minded students who initially enroll in MOOCs (Gannes, 2014).

Therefore, under the assumption that MOC dropout rates would be (say) no more than 50%, the economics of MOCs are compelling enough to challenge the current business model used by most universities (the majority of matriculated students take courses on-campus). In particular, at least the first and second year courses required for an undergraduate degree could be delivered by far fewer campuses using far fewer faculty and staff to much larger numbers of students. For instance, the largest four-year public university system, California State University (CSU) (Office of Public Affairs, 2013) offered 79,319 classes during their Fall, 2011 quarter. Of these, 27,524 were lower division undergraduate classes or 34.7% of the total (Office of the Chancellor, 2013). On average, a tenured or tenure-track faculty member of CSU teaches three classes per quarter (Office of the Chancellor, 2013). Therefore, if CSU were to offer all of their lower division classes as MOCs, the system would need at most one third less faculty. Colleges with more or less graduate course offerings than CSU would realize less or more savings. Of course this projection assumes that students are not offered an on-campus alternative to the MOC offering.

## **The Need for Tutoring**

Matriculated students taking a course for credit are in-part paying for support services such as tutoring. This becomes a major impediment to the cost-effective offering of MOCs because it is cost-prohibitive to offer traditional tutoring support for a class size of about 500 or more. This is because tutoring is typically provided through a tutor pool – about one tutor for every 32 students (Oxford, 2013) or ideally, one tutor for every 20-30 students (Lentell and O'Rourke, 2004). There is some evidence (Gannes, 2014) that because college credits are being pursued, students enrolled in a MOC may be more motivated to complete the course than they would be if they were taking the course as a MOOC. And a MOOC can easily have an enrollment of 1K and larger. To the extent that MOCs would enjoy such high enrollments, the need to provide tutoring then, makes the offering of a MOC infeasible under current tutoring support delivery systems.

To address this need for tutoring support, an *online intelligent tutoring system* (OITS) is proposed that, being artificially intelligent, can handle thousands of simultaneous tutoring sessions when run in an appropriately-sized cluster computing environment.

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/automating-a-massive-online-course-with-cluster-computing/222311](http://www.igi-global.com/chapter/automating-a-massive-online-course-with-cluster-computing/222311)

## Related Content

---

### An Adaptive Model for a Rigorous Professional Practice Doctorate: The Disquisition

Robert Crow, Kofi Lomotey and Kathleen Topolka-Jorissen (2016). *Contemporary Approaches to Dissertation Development and Research Methods* (pp. 205-220).

[www.irma-international.org/chapter/an-adaptive-model-for-a-rigorous-professional-practice-doctorate/156939](http://www.irma-international.org/chapter/an-adaptive-model-for-a-rigorous-professional-practice-doctorate/156939)

### How Can Human Resource Management Help the Theory of Constraints

Brian J. Galli (2019). *International Journal of Strategic Engineering* (pp. 1-13).

[www.irma-international.org/article/how-can-human-resource-management-help-the-theory-of-constraints/219320](http://www.irma-international.org/article/how-can-human-resource-management-help-the-theory-of-constraints/219320)

### A Hybrid Technique Using PCA and Wavelets in Network Traffic Anomaly Detection

Stevan Novakov, Chung-Horng Lung, Ioannis Lambadaris and Nabil Seddigh (2015). *Research Methods: Concepts, Methodologies, Tools, and Applications* (pp. 758-786).

[www.irma-international.org/chapter/a-hybrid-technique-using-pca-and-wavelets-in-network-traffic-anomaly-detection/124526](http://www.irma-international.org/chapter/a-hybrid-technique-using-pca-and-wavelets-in-network-traffic-anomaly-detection/124526)

### Dissemination of Scholarly Information Among the Research Scholars Through Social Networks and Media in Madurai Kamaraj University and Manonmaniam Sun

P. Pitchaipandi (2020). *Measuring and Implementing Altmetrics in Library and Information Science Research* (pp. 132-144).

[www.irma-international.org/chapter/dissemination-of-scholarly-information-among-the-research-scholars-through-social-networks-and-media-in-madurai-kamaraj-university-and-manonmaniam-sun/247751](http://www.irma-international.org/chapter/dissemination-of-scholarly-information-among-the-research-scholars-through-social-networks-and-media-in-madurai-kamaraj-university-and-manonmaniam-sun/247751)

### Application of Statistical Analysis Tools and Concepts to Big Data and Predictive Analytics to New Product Development

Brian J. Galli (2020). *International Journal of Strategic Engineering* (pp. 17-35).

[www.irma-international.org/article/application-of-statistical-analysis-tools-and-concepts-to-big-data-and-predictive-analytics-to-new-product-development/243666](http://www.irma-international.org/article/application-of-statistical-analysis-tools-and-concepts-to-big-data-and-predictive-analytics-to-new-product-development/243666)