Course Instructor:
Dr. Aravind Mohan
Office Location: Alden Hall 106
Office Phone: +1 8143322883
Email: amohan@allegheny.edu
Web Site: http://www.cs.allegheny.edu/sites/amohan/

Instructors Office Hours
- Monday, Wednesday, and Friday: 9:30 am –11:00 am (15 minute time slots)
- Tuesday: 11:00 am –12:00 noon (15 minute time slots)
- Thursday: 11:00 am –11:30 am (15 minute time slots)

To schedule a meeting with me during my office hours, please visit my website and click the Schedule link in the top right-hand corner. Now, you can browse my office hours or schedule an appointment by clicking the correct link and then reserving an open time slot. Students are also encouraged to post appropriate questions to a channel in Slack, which is available at https://CMPSC441s2018.slack.com/, and monitored by the instructor.

Course Meeting Schedule
Lecture: Tuesday and Thursday, 9:30 am –10:45 am at Alden 101
Lab: Monday, 1:30 pm –3:20 pm at Alden 101

Academic Bulletin Course Description
A study of the principles used in the design, implementation, and evaluation of distributed systems. Participating in hands-on activities that often require teamwork, students investigate topics such as remote communication, process synchronization, and fault tolerance while creating and assessing components of a distributed system. During a weekly laboratory session students use advanced distributed programming environments to complete projects, reporting on their results through both written reports and oral presentations.

Prerequisites: Computer Science 280 officially (but you may take it if you have the instructor permission)

Required Texts and Materials
There is not a single required textbook for this class. Reading assignments will be provided through Bitbucket repository and/or course webpage from several sources such as: books, research articles from a wide variety of conference proceedings, journals, and the popular press.

Learning Objectives
Upon successful completion of this class, the student will be able to:

1. Describe clearly the basic concepts of distributed systems and know how to use the Cloud to scale the setup for distributed systems.
2. Implement data collection modules to collect data.
3. Integrate data storage platforms through calling APIs.
4. Build distributed workflow driven approach to process big data using the Cloud.
5. Familiarize in using scalable techniques on existing distributed systems in the NoSQL world.
6. Understand and present scholarly work in the area of big data and cloud computing.
7. Develop stand-alone big data application using a NoSQL-based distributed systems.

Teaching and Learning Methods
The main mode of learning in this class is following along with the posted course material, completing lab work as instructed by the lab specifications, and reading the accompanying materials provided by the instructor. Students are responsible for reading online resources as needed to expand on the topics that are discussed in the lectures. The
instructor will ask questions to stimulate thinking and participation. Students comments and questions are highly encouraged during the class and via the course Slack channel.

Students are encouraged to form a team and interact with the instructor on brainstorming and developing new ideas for technically sound final project. The key for success is to identify the project idea as early as possible. It is also possible to design the final project as a stepping stone to build collaborations with the instructor (for example: independent research study, summer research project, and thesis work).

Students are responsible for attending each lecture and lab session when scheduled (see the Attendance Policy for further details). Course content will be delivered via BitBucket, and assignments should also be submitted to BitBucket.

**Grading and Evaluation**

Your total grade for the course will be based on the following, weighted appropriately:

- Quiz (20%)
- Labs (20%)
- Project (30%)
- Paper presentation (20%)
- Attendance and class participation (10%)

A more detailed breakdown of the expectations for grades in the course is as follows:

- **Class Participation:** All students are required to actively participate during all of the class sessions. Your participation will take forms of completing class exercises (individually and in teams), answering questions about the required reading assignments, contributing to class discussions and asking constructive questions. You must also regularly participate in the discussions and polls on the Slack channels for this course.

- **Labs:** The laboratory assignments invites students to experiment with various tools and techniques for designing, implementing, evaluating, and documenting your distributed system for processing big data at scale. To best ensure that students are ready to develop software after graduation, students will complete most of the lab assignments in teams. Many of the lab assignments in this course will expect students to give both a presentation and a demonstration of the system that they created.

- **Paper presentation:** Research paper presentation will allow students to survey the state-of-the-art research on the topic of distributed systems, big data, workflows, and cloud computing. Students share findings from the readings with their colleagues in the course.

- **Quizzes:** The quizzes will cover all of the material in its associated module. The date for each quiz will be announced at least one week in advance of the scheduled date, tentatively (three to five) quizzes are planned. Unless prior arrangements are made with the instructor, all students will be expected to take the quizzes on the scheduled dates.

- **Final Project:** This project will present students with an opportunity to design and implement a distributed system that process big data using the Cloud computing framework. Completion of the final project will require you to apply all of the knowledge and skills that you have acquired during the course of the semester.

**Assignment Submission and Late Policy**

Every assignment has a due date and time. Failure to hand in the assignment by the deadline will result in a late submission penalty.

Assignments handed in within one week of the deadline will receive automatic grade reductions of 20% (in addition to any points deducted for errors). Assignments will not be accepted more than one week past the deadline, unless you can provide documented extenuating circumstances. Any extenuating circumstances must be documented through the Learning Commons, Counseling Center, Dean of Students office, Health Center, or other authoritative source.

If you are unable to attend class or lab for any reason beyond illness or injury, you must make arrangements with me to turn in assignments before class. Presentations must be delivered at scheduled times. This includes the paper presentations and final project presentation. Please check with the instructor one week before making any travel plans.
for the end of the semester or around breaks.

Attendance Policy
It is mandatory for all students to attend all of the class and laboratory sessions. If you will not be able to attend a
session, then please see the course instructor at least one week in advance to describe your situation. Students who
miss more than five unexcused classes, laboratory sessions, or group project meetings will have their final grade in
the course reduced by one letter grade. Students who miss more than ten of the aforementioned events will
automatically fail the course.

Disability Statement
Students with disabilities who believe they may need accommodations in this class are encouraged to contact Student
Disability Services (SDS) at (814) 332-2898. SDS is part of the Learning Commons and is located in Pelletier
Library. Please do this as soon as possible to ensure that such accommodations are implemented in a timely fashion.

Email and Slack
The instructor will primarily be checking the course Slack channel and his allegheny email account on regular basis.
In general, you could expect the instructor to reply to your email messages during:

- scheduled office hours
- morning time between 8.00 am – 9.00 am
- afternoon time between 1.00 pm – 2.00 pm

The instructor does not usually check his email and slack during weekends. Hence, plan it accordingly to send an
e-mail to the instructor during week days. Students who are struggling with the course material or who have question
should begin by posting their question (unless a private concern) to the Slack channel, so that the instructor or a
fellow student can provide an answer within the bounds of the Honor Code.

Class Preparation
In order to minimize confusion and maximize learning, students must invest time to prepare for the class discussions
and lectures. During the class periods, the course instructor will often pose demanding questions that could require
group discussion, the creation of a program or data set, a vote on a thought-provoking issue, or a group presentation.
Only students who have prepared for class by reading the assigned material and reviewing the current assignments
will be able to effectively participate in these discussions. More importantly, only prepared students will be able to
acquire the knowledge and skills that are needed to be successful in both this course and the field of computer science.

Honor Code
All students enrolled at Allegheny College are bound by the Honor Code. It is expected that your behavior will reflect
that commitment. To this end, we expect that you will adhere to the following Department Policy:

Department of Computer Science Honor Code Policy

It is recognized that an important part of the learning process in any course, and particularly in computer science,
derives from thoughtful discussions with teachers, student assistants, and fellow students. Such dialogue is
encouraged. However, it is necessary to distinguish carefully between the student who discusses the principles
underlying a problem with others, and the student who produces assignments that are identical to, or merely
variations on, someone else's work. It will therefore be understood that all assignments submitted to faculty of the
Department of Computer Science are to be the original work of the student submitting the assignment, and should be
signed in accordance with the provisions of the Honor Code. Appropriate action will be taken when assignments give
evidence that they were derived from the work of others.

You are encouraged to periodically review the specifics of the Honor Code as stated in the College Catalogue, The
Compass, and elsewhere.

Additionally, the Honor Committee co-chairs have requested that a signature as well as the following phrasing be
included on all submissions of graded work:

"This work is mine unless otherwise cited."