May 18th, 2017

Large-scale Graph Processing and Applications

Shawn Zengxiang Li
Distributed Computing Group, A*Star Institute of High Performance Computing, Singapore

Abstract

In the Big Data era, graph processing has been widely used to represent complex system structure, capture data dependency and uncover relationship insights. Due to the ever-growing graph scale and algorithm complexity, several distributed graph processing frameworks have attracted many interests from both academia and industry. In this talk, I will investigate how to achieve the trade-off between performance and cost for large scale graph processing on the Cloud. System-aware and machine learning models are developed to predict the performance of distributed graph processing tasks. Consequently, cost-efficient resource provisioning strategies could be recommended by selecting a certain number of VMs with specified capability subject to the predefined resource price and user preference. At the end of this talk, I will briefly introduce our recent projects on urban computing, disease simulation and social network analytics based on graph processing and real world data.

Short bio

Dr. Shawn Zengxiang Li is a scientist in Institute of High Performance Computing, A*Star, Singapore. His research interests include big-data analytics, large-scale graph processing, simulation & modelling, parallel &

distributed computing, data center and Cloud computing. He is Principle Investigator and key member of several successful research and industry projects, which are relevant to resource management of multi-tenant data center, distributed data/graph processing frameworks, urban computing, transportation optimization, disease simulations, etc. He has published dozens of high quality papers on ACM/IEEE Transaction Journals and Conferences. Dr. Li served as program committee member of several International conferences, including PADS, DS-RT, CloudCom and many others.