"A walk on the intersection of systems for large scale data analytics, and data analytics for systems"

Χρήστος Κοζανίτης
Research collaborator
FORTH-ICS

Περίληψη – Abstract

Despite incredible recent advances in data analytics, building data-driven applications remains prohibitively time-consuming and expensive for all but the best-trained, best-funded engineering organizations. For instance, data collection, extraction, and cleaning requires tedious trial-and-error based software development. Moreover, Machine Learning model training requires the use of expensive infrastructure to run for days. And in the lack of better hardware, the deployment of models in production does not always use the best possible platforms.

For pipelines with completely different compute requirements between data acquisition, and statistical model fitting, my work uses proper abstractions that modify Apache Spark to work with clusters of heterogeneous executors. In terms of infrastructure costs, my work uses Machine Learning techniques to accommodate more workloads over the same compute cluster. And in order to achieve low latency speeds, which are crucial for Large Scale streaming engines, such as Apache Flink, we have been working on hardware implementations for Neural Network inference that dramatically reduce the latency and power consumption of predictions over Convolutional Neural Networks.

Christos Kozanitis is a research collaborator at FORTH-ICS. He received his M.S. and Ph.D in Computer Science and Engineering from the University of California, San Diego in 2009 and 2013 respectively. Parts of his phd work influenced products from companies such as Cisco and Illumina. He also held a two-year postdoctoral appointment at the AMP Lab of the University of California, Berkeley, where he used and adapted state of the art big data technologies, such as Apache Spark SQL, Apache Parquet and Apache Avro to process large amounts of DNA sequencing data. His current research interests involve the improvement in software, storage and hardware level of modern datacenters in order to speed up the processing of big data workloads.

ΔΙΑΛΕΞΗ

Παρασκευή 28/6/2019 – 12:00
Αίθουσα Σεμιναρίων,
Κτίριο Μηχανικών Η/Υ & Πληροφορικής
Πανεπιστήμio Ιωαννίνων