

Taiwan NCU Lecture

Title: Optimizations in extreme intensive data computation

Abstract: At the edge of the extreme data intensive computation, computer scientists and engineers are facing many challenges. As Thom Dunning had advocate, the ever increasing amount of data is rushing toward the exascale while we are just deploying petascale facilities such as 'Titan' or 'Blue Waters'. Improvements in hardware, from storage to processing unit and networks needed to handle this unparalleled evolution, will not be sufficient and need to rethink tools to describe and execute heavy computations.

Computation is a matter of knowledge representation, from source code to compilation, execution of the code to visualization of results. As big data emphasizes every step from classical programs it let room for new kind of optimization and new tools. As in many parts of computer innovation, automation and learning can lead to substantial improvements in resource usage. Among the many leads, we are here starting to use data mining and machine learning to define possible optimizations in issue from choice of data structures to resource utilization.

During this lecture, Nicolas Kamenoff will advocate for a contribution of programmers and computer scientists to face the data issue. He will also explain the vision of big data behind the starting ICoDF project (Intensive Computation Data Framework), that encompass best practices in knowledge representation and artificial intelligence to optimize many steps of big data solutions, and will present the next steps of this project related to the TWEA-DC (Taiwan Extragalactic Astronomical Data Center).