
Reliability Issues in Data Intensive Computing

Mootaz Elnozahy
(IBM)

Abstract

The continuing growth in data size and processing requirements is challenging today's computing capabilities for a class of applications that we have come to classify as data-intensive computations. In this talk, we focus on the reliability aspects of this challenge. We argue that traditional reliability techniques based on space or time redundancy are reaching their limits for these applications. The growth in data size is making it expensive to rely on replication techniques. Furthermore, the power consumption limitations on today and future systems will limit the possibilities of relying on process replication or rollback-recovery. Therefore, we explore the possibilities of using a new framework based on approximate and/or imprecise computations. In this framework, accuracy is traded off for reduced power consumption and improved performance. This work is still nascent, so we will focus on the problems figuring out how to draw the dividing line between acceptable and unacceptable results, and to identify the division of responsibilities between system software, application software and hardware in carrying out such computations.
