A Survey on Data-Centric and Data-Aware Techniques for Large Scale Infrastructures

Authors : Silvina Caíno-Lores, Jesús Carretero

Abstract : Large scale computing infrastructures have been widely developed with the core objective of providing a suitable platform for high-performance and high-throughput computing. These systems are designed to support resource-intensive and complex applications, which can be found in many scientific and industrial areas. Currently, large scale data-intensive applications are hindered by the high latencies that result from the access to vastly distributed data. Recent works have suggested that improving data locality is key to move towards exascale infrastructures efficiently, as solutions to this problem aim to reduce the bandwidth consumed in data transfers, and the overheads that arise from them. There are several techniques that attempt to move computations closer to the data. In this survey we analyse the different mechanisms that have been proposed to provide data locality for large scale high-performance and high-throughput systems. This survey intends to assist scientific computing community in understanding the various technical aspects and strategies that have been reported in recent literature regarding data locality. As a result, we present an overview of locality-oriented techniques, which are grouped in four main categories: application development, task scheduling, in-memory computing and storage platforms. Finally, the authors include a discussion on future research lines and synergies among the former techniques.
Keywords : data locality, data-centric computing, large scale infrastructures, cloud computing
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020