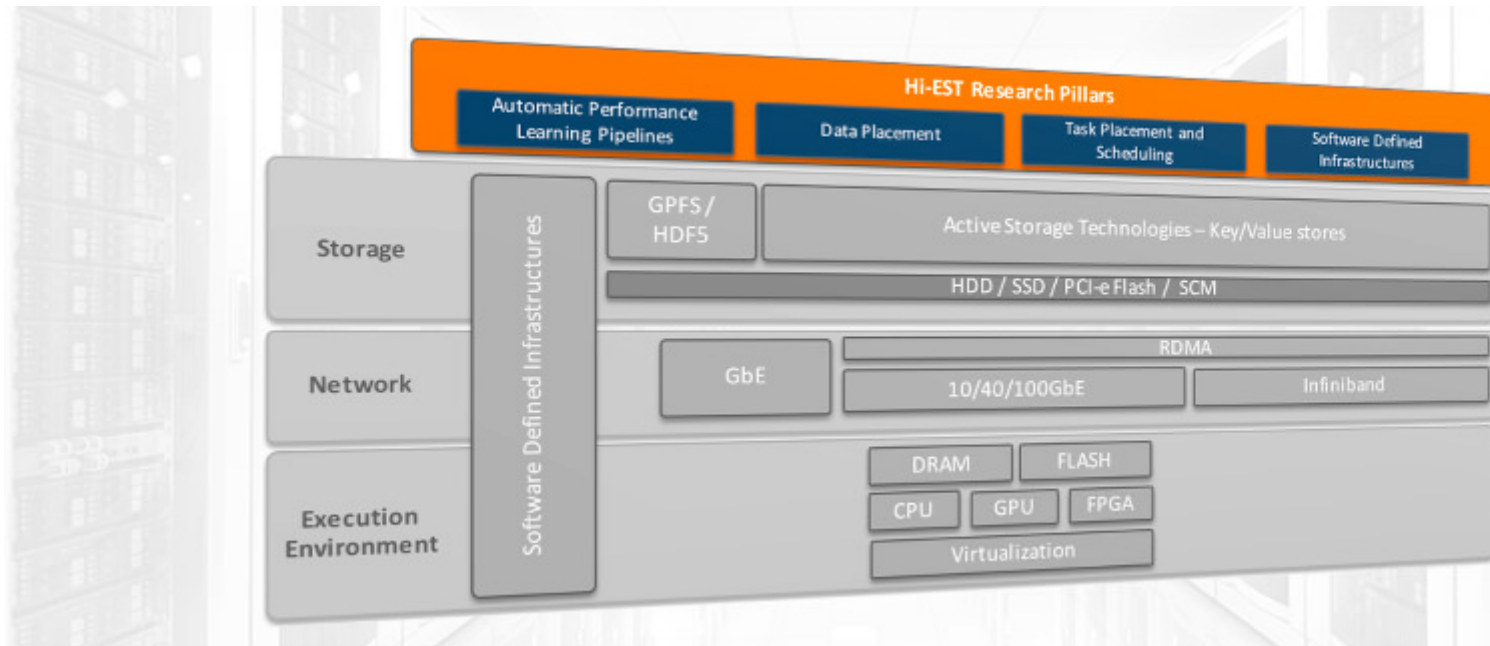


Data-Center Optimization



This research line focuses on the development of data center optimization methods for the upcoming Software Defined Infrastructures, leveraging learning techniques to build models that guide optimization techniques

Summary

The complexity of the data-center scenarios that start emerging is going to be extremely high. The set of technologies and workloads that are expected to be hosted by future data-centers challenges all currently existing technologies in terms of resource management. In the past, heuristics have been used as workarounds to address NP-Hard placement problems with variable results. The major problem with such heuristics is that they need to make generic assumptions on some workload, user or hardware characteristics that may hardly adapt to changing conditions.

Portability of these heuristics across facilities and domains is also limited, as practical approaches that seem to fit very well to the characteristics of a particular environment can easily be insufficient when the scenario is changed. For this reason, in we approach the problem of managing self-knowledge from the point of view of automatized learning techniques. Machine Learning has multiple applications in modern data centre workloads. In this research line, these techniques will be taken one step forward to capture performance properties of different workloads, and will be designed and trained to extract the essence of existing relations between workloads, technologies and performance.

Objectives

The goal of this research line is to perform a significant advance in the field of methods, mechanisms and algorithms for the integrated management of heterogeneous workloads in software defined infrastructures. In particular, the project aims to achieve the following objectives:

- Advance research frontiers in Adaptive Learning Algorithms by proposing the first known use of Deep Learning techniques for guiding task and data placement decisions
- Advance research frontiers in Task Placement and Scheduling by studying the first known algorithm to map heterogeneous sets of tasks on top of systems enabled with Active Storage capabilities
- Advance research frontiers in Data Placement strategies by studying the first known algorithm used to map data on top of heterogeneous sets of key/value stores connected to Active Storage technologies
- Advance research frontiers in Software Defined Environments by developing the still inexistent vocabulary to describe Supercomputing workloads and an associated language

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 13 jul 2024 - 21:06): <https://www.bsc.es/ca/research-development/research-areas/cloud-computing/data-center-optimization>