Enabling Extreme-scale science with RADICAL-Cybertools

George Koubbe
g.koubbe@rutgers.edu

Kartik Rattan
kr642@scarletmail.rutgers.edu

Advisor: Prof. Shantenu Jha, Dr. Matteo Turilli
http://radical.rutgers.edu

RADICAL Research

The **R**esearch in **A**dvanced **DI**stributed **C**yberinfrastructure and **A**pplications **L**aboratory (RADICAL) works at the intersection of Computing, Computational & Data-enabled Science and Cyberinfrastructure Research & Development.

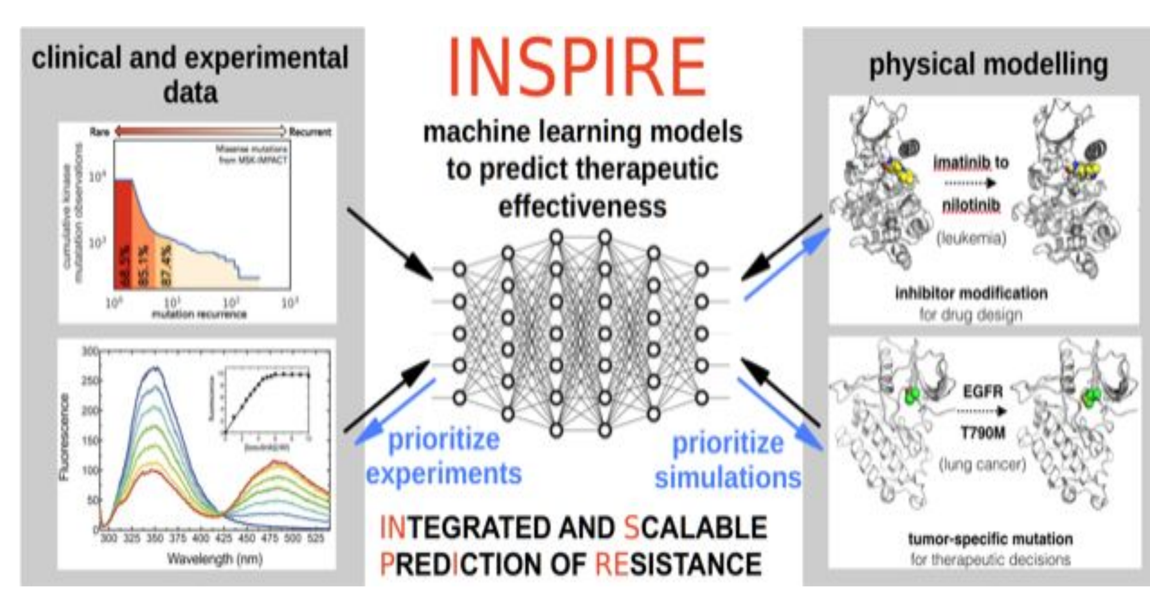
The RADICAL team uses high-performance distributed computing to advance diverse science and engineering problems at scale, in ways that would not be possible without RADICAL solutions. The RADICAL team has enhanced research in several domains, from biomolecular sciences to high energy physics and polar science. We have designed and developed RADICAL-Cybertools, which is described next.

You can find currently active and past research projects the Lab was involved with: http://radical.rutgers.edu/projects.

RADICAL-Cybertools Overview

RADICAL-Cybertools (RCT) are an abstractions-based suite of software modules designed to support scalable, interoperable and sustainable science on a range of high-performance and distributed computing infrastructure. The four main RCT modules are: RADICAL-SAGA, RADICAL-Pilot, RADICAL Ensemble Toolkit, and RADICAL-Analytics.

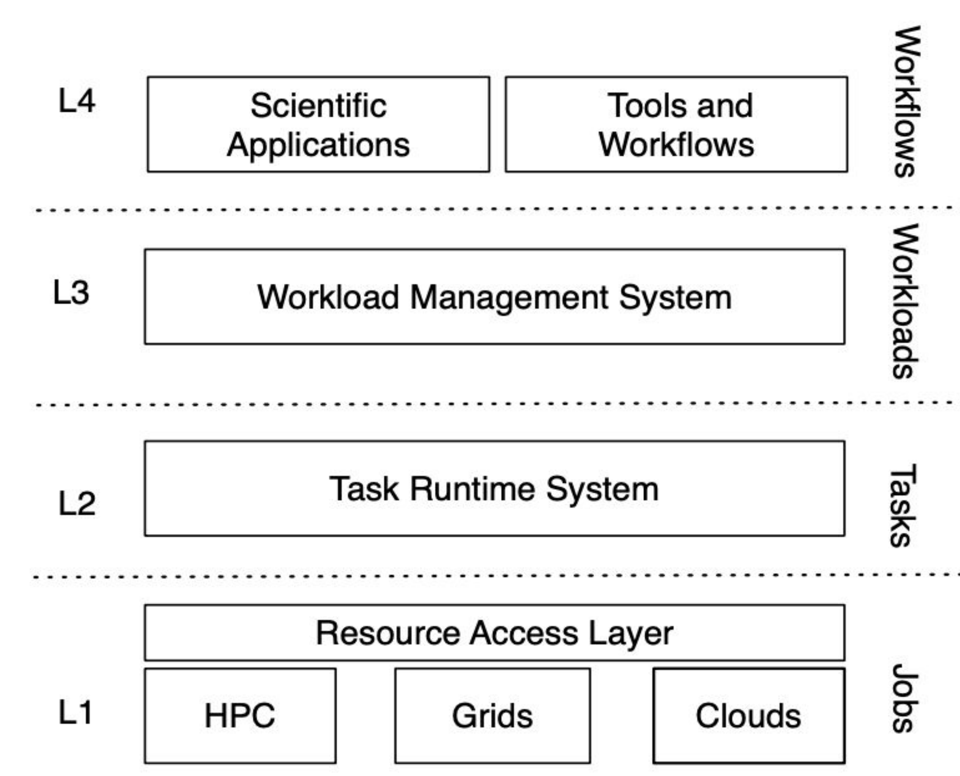
RCT enables the utilization of the scale and power of supercomputers, to facilitate scientific discovery in a range of domains: climate, earth and biomedical sciences.



RADICAL-Cybertools Design

RCT are designed and implemented in accordance with the building blocks approach: self-sufficient, interoperable, composable, and extensibile.

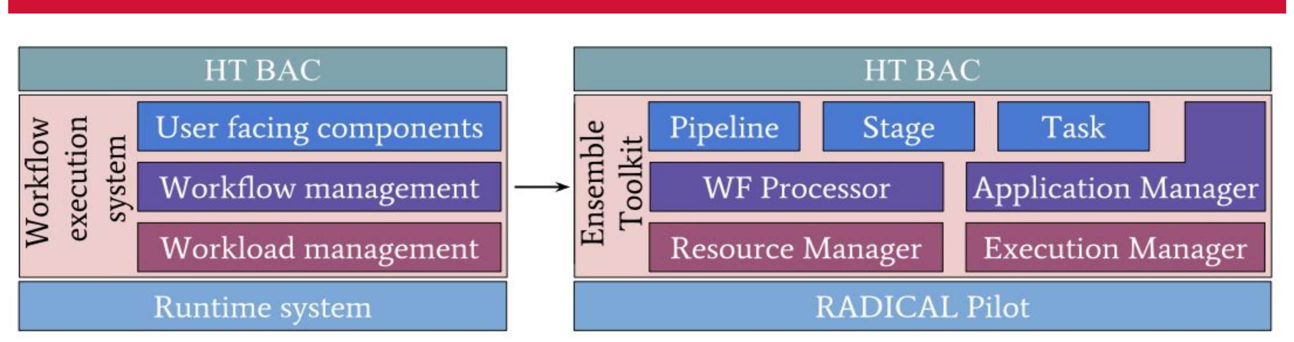
L4 Workflow and Application Description: Requirements and semantics of an application described in terms of a workflow.
L3 Workload Management System: Responsible for managing and configuring the available resources for the given workload.
L2 Task Runtime System : Responsible for effective and efficient task execution on the selected resources.
L1 Resources: Resources used to execute the tasks characterized by their capabilities, availability and interfaces.



ICEBERG

- **Seals**: Automate pack-ice seal surveys using a combination of remote sensing, computer vision, seal ecology and HPC.
- **Penguins**: A detection algorithm to extract the location of penguin colonies from high-resolution imagery.
- **Rivers**: Detect regions of large and small rivers and to distinguish them from crevasses and non-channelized slush
- **LandCover**: Automated processing of satellite imagery, i.e., removal of snow, ice, water, and shadows from the scene.

Extreme Scale Binding Free Energy Calculations



Challenge: Rapid and accurate quantification of the molecules binding affinity to their target proteins in order to improve drug design and treatment, providing patient-specific medicine in cases like cancer therapy.

Solution: Adaptive approaches on ensemble-free energy protocols with the goal of reducing computational resources and improving efficiency.

Among these protocols, we have ESMACS and TIES, both designed to use an ensemble MD simulation, with a focus on intra-protocol adaptivity which relies on intermediate runtime results within a protocol instance to define the following set of simulations. For this: **HTBAC**.

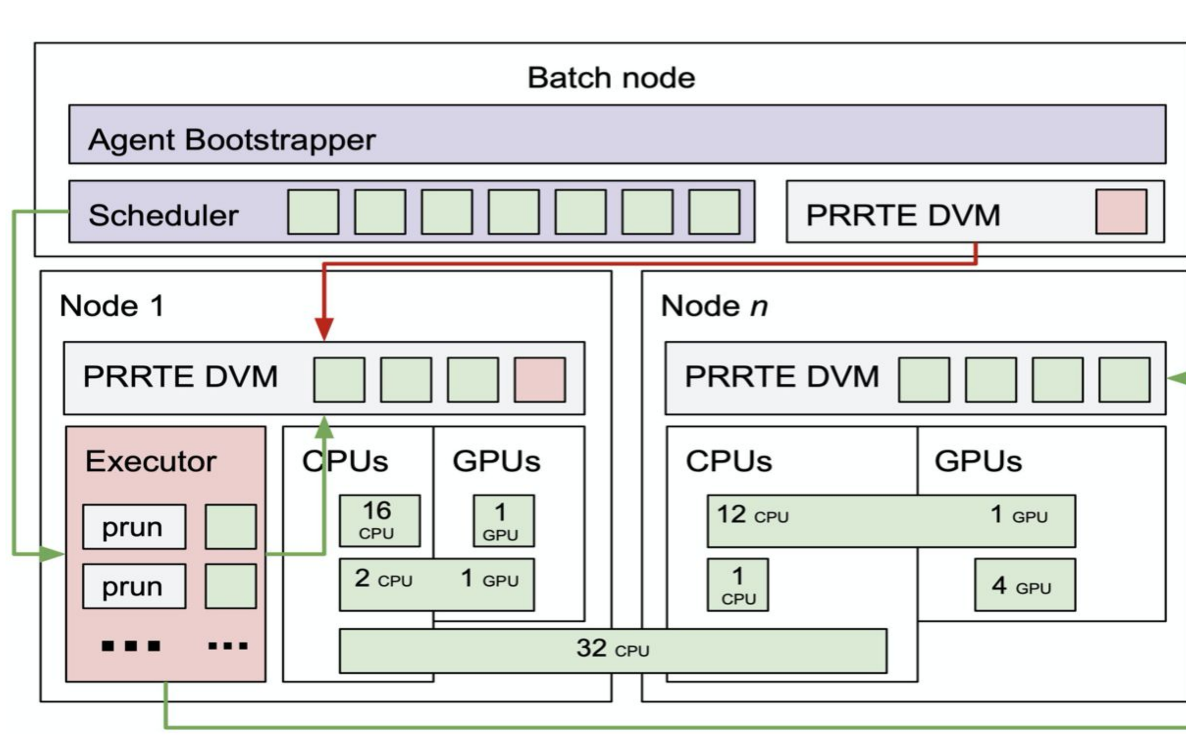
HTBAC makes use of RADICAL's EnTK and RP, and it is implemented in Python as a domain-specific library.

Many-tasks applications on High Performance Computing platforms

Workloads of many scientific applications are comprised of many tasks:

- Tasks are independent and self contained
- Tasks can be heterogeneous
- HPC adoption is required for both scale and performance ==> challenging!

To address this, we have characterized the performance of executing many tasks using **RADICAL-Pilot** when interfaced with JSM and PRRTE on Summit.



Conclusion: RP can effectively integrate with both JSM and PRRTE, imposing manageable aggregated overheads while offering high degrees of configurability.

Acknowledgement

In addition to support from Rutgers we acknowledge financial support of NSF 1443054.

For more information about our projects scan the QR code