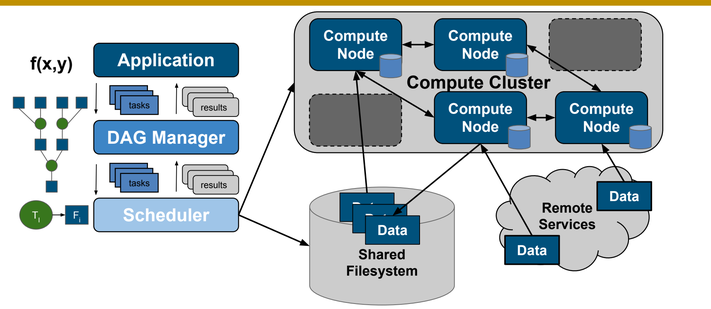
Reshaping High Energy Physics Applications for Near-Interactive Execution Using TaskVine

Barry Sly-Delgado(bslydelg@nd.edu) - University of Notre Dame - http://ccl.cse.nd.edu/

Abstract

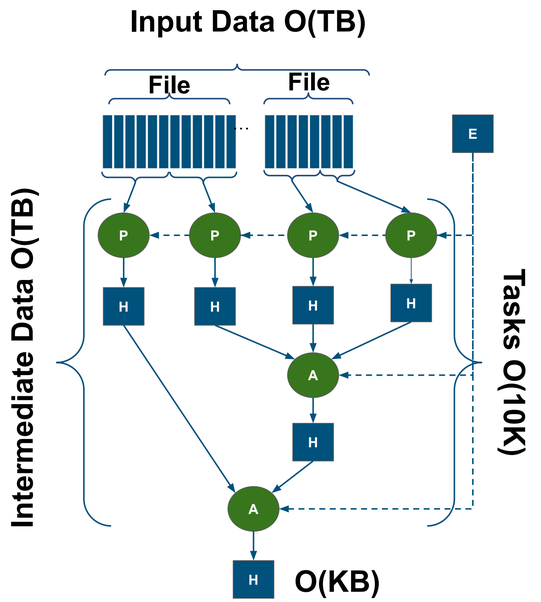
The early stages of high energy physics analyses involve processing large sums of data using long-running distributed workflows to reduce the initial amount of data into manageable datasets, often taking weeks. In turn, the latter parts of an analysis can take hours to produce results. However, latter analysis applications can be reshaped, resulting in near-interactive execution times. To facilitate application reshaping, improvements to application stack must be made. Improvements to the hardware capabilities of a compute cluster are one avenue to reduce the overall runtime of a given application by improving data access speeds for initial datasets. Additionally, improvements to the task scheduler can greatly reduce runtime by reducing data movement through improved data management techniques. Additionally, improved task execution paradigms reduce the overhead for task startup. In this paper, we present the improvements made via TaskVine, a workflow scheduler, that facilitated application reshaping for real-world physics application DV3.

Application Stack

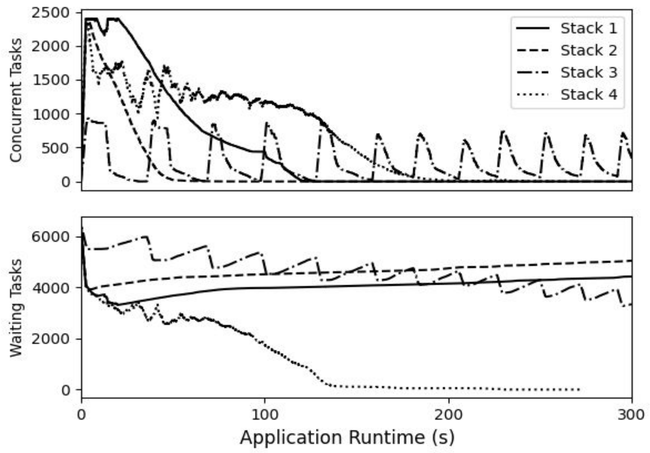


	Stack 1	Stack 2	Stack 3	Stack 4
Application	DV3 / Coffea	DV3 / Coffea	DV3 / Coffea	DV3 / Coffea
DAG Mgr.	Dask	Dask	Dask	Dask
Scheduler	Work Queue	Work Queue	TaskVine +PythonTask	TaskVine + Functions
Storage	Hadoop	VAST	VAST	VAST

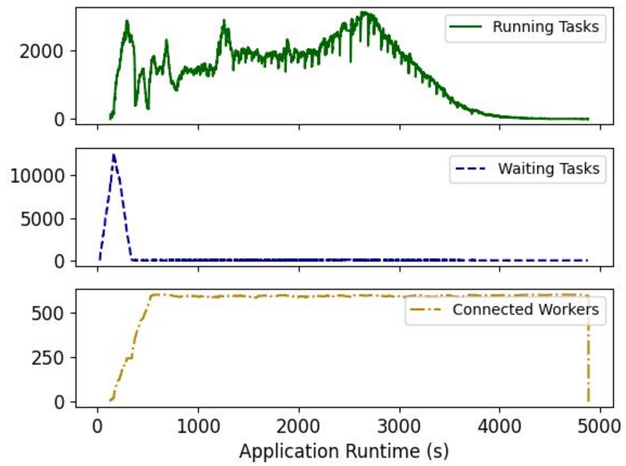
HEP Analysis Workflow



Improvements

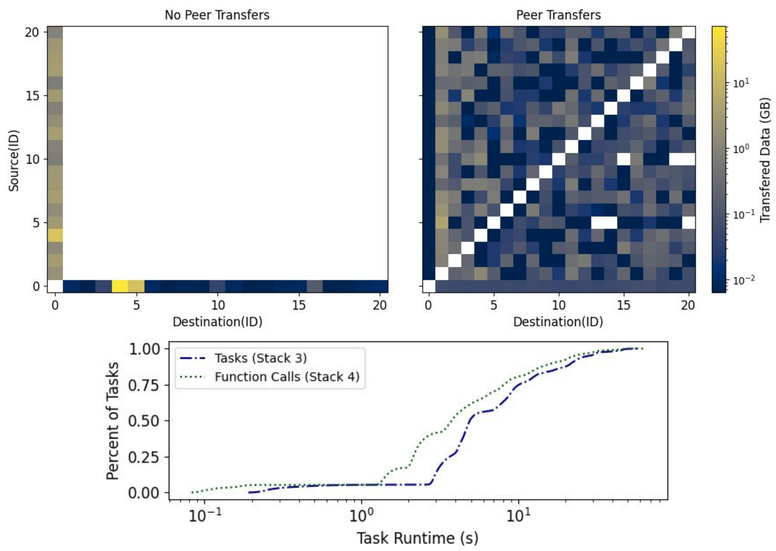


(Above) DV3 application executed on stacks 1, 2, 3, and 4. This configuration contains roughly 7500 tasks. Each run executed on 200 12-core workers.



(Above) Full scale execution of DV3. This workflow contains 175,000 tasks, executed on 600 12-core workers.

Stack Progression



(Left Top) - Execution of DV3 with and without peer transfers (stack 1 & 2 and 3 & 4 respectively).

(Left Bottom) - DV3 task runtimes with Tasks (stack 1, 2 snd, 3) and Function Calls/serverless (stack 4)

This work was supported in part by grant OAC #1931348
"CSSI Elements: Data Swarm: A User-Level Framework for Data Intensive Scientific Computing".

CCTools

UNIVERSITY OF NOTRE DAME

TaskVine