LLNL-JRNL-819205

LAWRENCE LIVERMORE NATIONAL LABORATORY

A three-phase workflow for general and expressive representations of nondeterminism in HPC applications

D. Chapp, D. Rorabugh, K. Sato, D. H. Ahn, M. Taufer

February 5, 2021

International Journal of High Performance Computing Applications

Scalable Composition and Analysis Techniques for Massive Scientific Workflows

Anonymous Author(s)

ACM Reference Format:
Anonymous Author(s). 2021. Scalable Composition and Analysis Techniques for Massive Scientific Workflows. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Current economic factors are ushering in a new era of extremely heterogeneous high performance computing (HPC). As Moorse's Law is tapering, specialized hardware such as graphics processing units (GPU) [12] is increasingly replacing the role that general purpose processors used to play as the main computing workhorse. As the cloud computing has become a dominant market force [6, 7], HPC has also begun to embrace a heterogeneous software environment, marrying its software stack with the cloud solutions. Indicatively, 2018 was the first time that new additions to the biggest 500 supercomputers derived more performance from specialized processors than from general purpose processors. The key cloud solutions including container and container orchestration technologies such as Kubernetes [9] and Red Hat OpenShift [**?**] systems are furthermore making their inroads into HPC infrastructure [3, 5]. HPC is undergoing an economic inflection point at which it must embrace the trends towards heterogeneity lest it lose its long-term viability

The required complexity of science workflows also grows sharply on HPC systems. The convergence of traditional HPC and new simulation, analysis, and data science approaches including machine learning (ML) and artificial intelligence (AI) provides unprecedented opportunities for discovery, but also creates workflows [1, 4, 10, 14] that are far more complex than traditional ones. This trend is further necessitated and accelerated by the extreme heterogeneity: Different workflow tasks [2, 4, 8] must be mapped to different specialized hardware (e.g., CPU, GPU, disaggregated AI accelerators [**? ?**], I/O storage of certain tier in a multi-tiered storage subsystem [13]) and/or different system software (e.g., persistent container services such as a message queue and database services running on on-premises Kubernetes cluster). The sharply increased complexity, however, has proven to be difficult for researchers to manage using the traditional monolithic, custom workflow approach.

One of the alternative workflow design architectures that significantly emerge amidst these trends is called *composite science workflows.* In this architecture, researchers select a set of preexisting composable workflow-management software components and flexibly blend them with a disparate set of domain-specific application and management software so as to create a scalable end-to-end science workflow. This approach is clearly gaining a traction for high-end computing. As a specifically recent demonstration, the winner [**?**] and two [8**?**] of three finalists of the SC20 Gordon Bell Special Award [1] for COVID-19 competition leveraged the scalable and composable workflow technologies [11].

Though on a promising path, attempts to realize the full potential of heterogeneous HPC centers via scalable composite science workflows is a burgeoning field. Generally speaking, the technologies employed for composite workflows have been developed in isolation and often feature widely varying levels of performance, scalability and interoperability. These factors make composing and optimizing the end-to-end workflow still highly daunting. Many workflow building blocks are emerging, and a myriad of analysis tools exist for each traditional HPC programming paradigm (e.g., a single application running at scale), there is a paucity of workflow techniques and tools that researchers can use to compose an end-to-end workflow on top of these blocks, and to analyze and to optimize its performance.

Figure 1 shows.

2 Motivational Example

MuMMI (Current job injection rate bottleneck?) Maybe linking that to Flux scheduling specialization's need for broader analysis??? ConveyorLC (task-level scheduler scalability and how it requires DAT – not optional to be used in automated workflow) Talk more about related work but not too extensive

3. Requirements for "composite science workflows" as exhibited by COVID-19 drug design workflow composition (Stephen, Jeff)

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

[1] Often called the Nobel Prize of supercomputing, the Association for Computing Machinery (ACM) Gordon Bell Prize is one of the most prestigious awards that recognizes outstanding achievement in HPC applications.

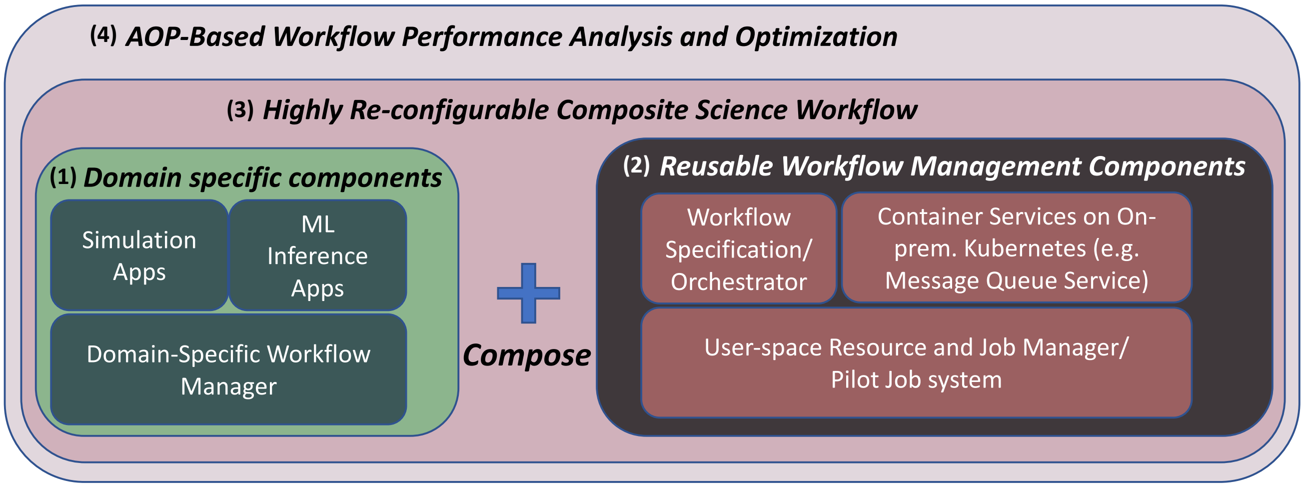

Figure 1. Easy-to-analyze and -optimize composite science workflow architecture

Briefly discuss the science goal of COVID-19 drug design workflow (Felice and Jonathan) Computation and workflow requirements (including k8s services)

4. Critical-path analysis driven scalable coupling techniques (Our approach) (Ahn)

Building blocks like Flux (Stephen) AOP and PerfFlowAspect (needs a way to cast cross-cutting performance concerns to be effective) (Ahn) Lessons learned and key required technical contributions for workflow optimization (Ahn propose key lessons first and then recruit the folks who can write each lesson the best)

5. Results

Description of systems (Lassen? Corona? Ruby?) and software (e.g., compiler, OS) used (Karlin) (Ahn want to spend one to two more weeks to determine exactly.) Individual components-wide evaluation ConveyerLC coupling analysis and Results (Before -> After) GMD scalability (Before -> After) Some microbenchmarks including Flux and other WF scheduling bottlenecks End to End workflow performance measurement and optimization Resulting composition AOP-based critical path information for workflow optimization If we can bring up and run the overall end to end in this time frame, shows a few key optimization Yes, really need a "case study" that removes bubbles in the pipeline and identifies critical components in which to focus efforts to optimize their specific performance (long or otherwise rate-limiting pipeline stages)

6. Related work and Discussion (Stephen/Frank)

7. Conclusion (Karlin and Brian) Try not to couch this as Future work but this lays out a foundation

Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-CONF-xxxxxx).

References

[1] Dong H. Ahn, Ned Bass, Albert Chu, Jim Garlick, Mark Grondona, Stephen Herbein, Helgi I. Ingólfsson, Joseph Koning, Tapasya Patki, Thomas R.W. Scogland, Becky Springmeyer, and Michela Taufer. 2020. Flux: Overcoming scheduling challenges for exascale workflows. *Future Generation Computer Systems* 110 (2020), 202–213. https://doi.org/10.1016/j.future.2020.04.006

[2] Harsh Bhatia, Francesco Di Natale, Joseph Y. Moon, Xiaohua Zhang, Joseph R. Chavez, Fikret Aydin, Chris Stanley, Tomas Oppelstrup, Chris Neale, Sara Kokkila Schumacher, Dong H. Ahn, Stephen Herbein, Timothy S. Carpenter, Sandrasegaram Gnanakaran, Peer-Timo Bremer, James N. Glosli, Felice C. Lightstone, and Helgi I. Ing. [n.d.]. Generalizable Coordination of Large Multiscale Workflow: Challenges and Learnings at Scale. In *to Appear in the Proceedings of Supercomputing '19: The International Conference for High Performance Computing (SC '21)*.

[3] National Energy Research Scientific Computing Center. [n.d.]. SPIN. https://www.nersc.gov/systems/spin/. Retrieved July 21, 2021.

[4] Francesco Di Natale, Harsh Bhatia, Timothy S. Carpenter, Chris Neale, Sara Kokkila Schumacher, Tomas Oppelstrup, Liam Stanton, Xiaohua Zhang, Shiv Sundram, Thomas R. W. Scogland, Gautham Dharuman, Michael P. Surh, Yue Yang, Claudia Misale, Lars Schneidenbach, Carlos Costa, Changhoan Kim, Bruce D'Amora, Sandrasegaram Gnanakaran, Dwight V. Nissley, Fred Streitz, Felice C. Lightstone, Peer-Timo Bremer, James N. Glosli, and Helgi I. Ingólfsson. 2019. A massively parallel infrastructure for adaptive multiscale simulations: modeling RAS initiation pathway for cancer. In *Proceedings of Supercomputing '19: The International Conference for High Performance Computing (SC '19)*. https://doi.org/10.1145/1122445.1122456

[5] David Fox. [n.d.]. Enabling Workflows in Livermore Computing. https://hpc.llnl.gov/sites/default/files/Enabling-Workflows-in-Livermore-Computing-LC-User-Meeting-Dec-2020%20Final.pdf. Retrieved July 21, 2021.

[6] Gartner, Inc. [n.d.]. Gartner forecasts worldwide public cloud end-user spending to grow 23% in 2021. https://www.gartner.com/en/newsroom/press-releases/2021-04-21-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-grow-23-percent-in-2021. Retrieved June 20, 2021.

[7] Hyperion Research. [n.d.]. How cloud computing is changing HPC spending. https://hyperionresearch.com/wp-content/uploads/2021/01/Hyperion-Research-Special-Analysis-Clouds-and-HPC-December-2020.pdf. Retrieved June 20, 2021.

[8] Sam Ade Jacobs, Tim Moon, Kevin McLoughlin, Derek Jones, David Hysom, Dong H Ahn, John Gyllenhaal, Pythagoras Watson, Felice C Lightstone, Jonathan E Allen, Ian Karlin, and Brian Van Essen. 2021. Enabling rapid COVID-19 small molecule drug design through scalable deep learning of generative models. *The International Journal of High Performance Computing Applications* (May 2021). https://doi.org/10.1177/10943420211010930 arXiv:https://doi.org/10.1177/10943420211010930

[9] Kubernetes. [n.d.]. Production-Grade Container Orchestration. https://kubernetes.io. Retrieved July 21, 2021.

[10] Amanda J. Minnich, Kevin McLoughlin, Margaret Tse, Jason Deng, Andrew Weber, Neha Murad, Benjamin D. Madej, Bharath Ramsundar, Tom Rush, Stacie Calad-Thomson, Jim Brase, and Jonathan E. Allen. 2020. AMPL: A Data-Driven Modeling Pipeline for Drug Discovery. *Journal of Chemical Information and Modeling* 60, 4 (2020), 1955–1968.

[11] Rob Farber. 2020. Workflow Technologies Impact SC20 Gordon Bell COVID-19 Award Winner and Two of the Three Finalists. https://www.exascaleproject.org/workflow-technologies-impact-sc20-gordon-bell-covid-19-award-winner-and-two-of-the-three-finalists.

[12] Neil C. Thompson and Svenja Spanuth. 2021. The Decline of Computers as a General Purpose Technology. *Commun. ACM* 64, 3 (Feb 2021), 64–72.

[13] Tiffany Trader. 2021. Livermore's El Capitan Supercomputer to Debut HPE 'Rabbit' Near Node Local Storage. https://www.hpcwire.com/2021/02/18/livermores-el-capitan-supercomputer-hpe-rabbit-storage-nodes.

[14] Chih-Chieh Yang, Giacomo Domeniconi, Leili Zhang, and Guojing Cong. 2020. Design of AI-Enhanced Drug Lead Optimization Workflow for HPC and Cloud. In *IEEE International Conference on Big Data*. 5861–5863.