Missing Abstractions

Name: Daniel S. Katz & Shantenu Jha

Summary: Without fundamental advances in how we think about distributed applications and systems that address heterogeneity and dynamism, we will remain condemned to point and non-extensible solutions

Related work: M. Turilli, F. Liu, Z. Zhang, A. Merzky, M. Wilde, J. Weissman, D. S. Katz, S. Jha, “Integrating Abstractions to Enhance the Execution of Distributed Applications,” IPDPS 2016, doi: 10.1109/IPDPS.2016.64.

Big/New Ideas of Missing Abstractions

We traditionally/mostly think of resources as static and homogenous

Resources are really **dynamic:** accessible, not accessible, fully available, partially available, etc.

.. and **heterogeneous**: compute speed, storage capacity, network latency and bandwidth to/from other systems

Lack of good abstractions

—> problems building software to use resources

e.g., coupling data transfer & storage w/ compute

—> unclear how to federate dynamic and heterogeneous resources

Can use simple resources, but can't really combine them

Current solutions work, but are not general or portable

Like 18th century medicine - we don't have the theory that explains why it works or what would happen in a slightly different situation

Indicated R&D for Missing Abstractions

Can we throw out our old ideas and start thinking fresh?

Parallel computing does not give right metaphor. What does?

How to separate models and implementations?

E.g., for resource management, execution, federation

Need to consider all relevant heterogeneous aspects simultaneously and at the same level (compute, network, storage)

Understand how to support different user interfaces, from individual programmers writing their own code to groups using community toolkits (ala ROOT) to groups using science gateways

Questions for the Software Institute

How do we generalize specific HEP (and non-HEP) solutions?

Do these help us determine/find “missing” abstractions?

Can we publicize them, then use them as a community?

… to build software that continues to work as computing platforms change

Meta-questions

What part of these questions can the institute answer?

What can it encourage others to do, then leverage?

How can HEP/HSF work with other large projects to find abstractions and build general tools/solutions?