Title: OpenHPC: a Collaborative Environment for Development of a Cohesive Comprehensive HPC Software Stack Suitable for Cloud Integration

Presenter: Robert W. Wisniewski

Over the last two decades there has been a trend toward requirements that are leading to an increasingly complex software stack on supercomputers. Supercomputers no longer sit in the corner of a data center, disconnected, running batch simulations. Instead, they have started to become an integral element of the data center and active participants in application workflows. While this trend has been going on for a while, in the last couple years it has significantly accelerated. Uncertainty quantification and multi-scale simulations that started this accelerated trend, led the way to more sophisticated analytics and the need for big data. More recently, to meet the needs of these applications and emerging ones, such as bio-informatics, machine learning, and others, ideas like extending the software stack to support containers, virtualizing the network, providing quality of service, and unifying services with cloud environments are being contemplated.

The effort to construct a software stack, while already significant, is going to increase substantially, and as noted above, there are interesting and new challenges that need to be solved. While there is a lot of open source work occurring in the open source community, these efforts are often independent and need to perform replicated work not relevant to the research path they are exploring. Finally, once an HPC researcher produces a new capability and is interested in a second group experimenting or leveraging it, depending on the infrastructure of the second site, the researcher may run into challenges because the software infrastructure at the second site may not be compatible with the stack at the researcher's site. What is needed is an open community focused on the HPC software stack that removes the replicated work and provides a cohesive environment facilitating collaboration.

In November 2015, over 30 organizations (has now grown to over 40 – see Figure 1), including vendors, OEMS, ISVs, universities, and HPC centers, announced the intent for form OpenHPC, an open source community focused on producing a coherent and comprehensive stack for HPC machines. The initial seeding of OpenHPC was an HPC software stack for mid-range (~1000 node) HPC use. That stack (see Figure 2) was purposely designed to be modular, allowing components to be replaced. This allows researchers or other organizations to leverage the breadth of the community in providing feedback on their component. In early discussions, community members have expressed an interest in contributing new components including ones geared toward big data and cloud, and that will be an early topic of the technical steering committee. For the type of work being described at BDEC, OpenHPC provides a favorable environment because it allows researchers to leverage the common infrastructure and more easily share the work they are doing. In this presentation, I will describe the vision behind OpenHPC, the status and progress since its intent to be created was announced, and describe some suggested contributions in the cloud space.

Figure 1: OpenHPC participation as of May 3, 2016

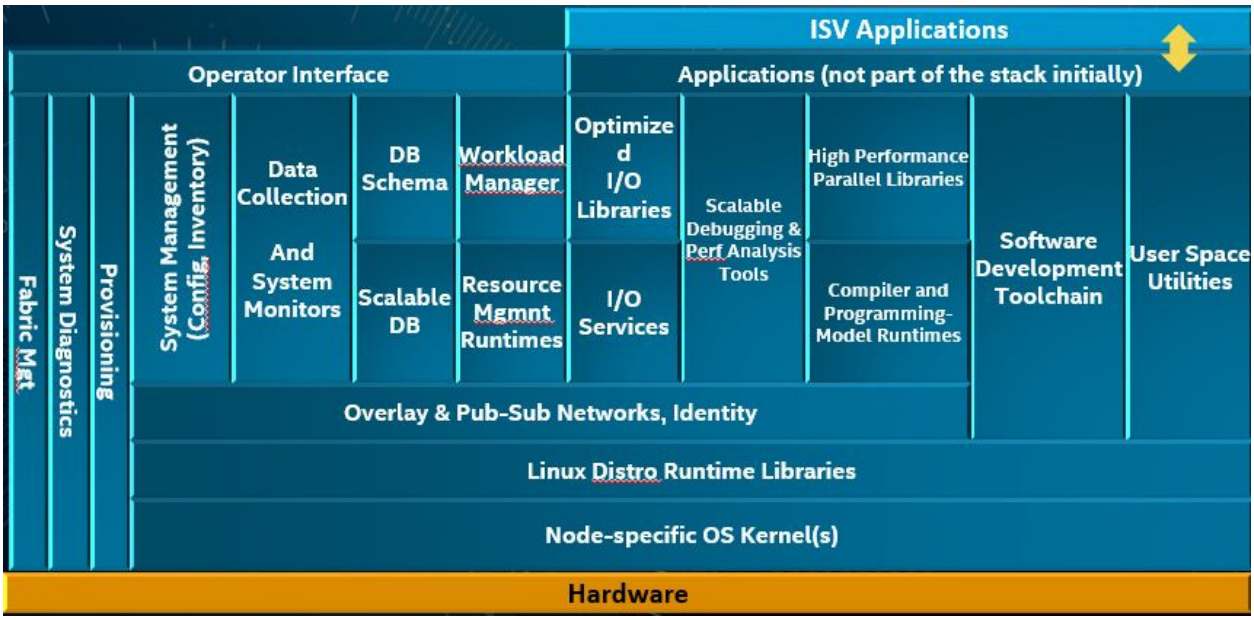

Figure 2: OpenHPC Modular Software Stack