

Emerging Computational Workloads Demand Virtual Clusters for Extreme-scale Systems

Andrew J. Younge, Kevin Pedretti, and Ron Brightwell

Center for Computing Research

Sandia National Laboratories

P.O. Box 5800, MS-1319

Albuquerque, NM 87185-1110

Email: {ajyounge}@sandia.gov

Abstract—Large Scale Data Analytics are gaining attention not only as a processing component to larger simulation workloads, but also as standalone scientific tools for knowledge discovery, yet system software for such capabilities on the latest extreme-scale DOE systems has failed to support these types of application ecosystems. Instead of adapting these workloads to traditional batch scheduling systems, we need a way to provision Virtual Clusters on advanced supercomputing resources to enable diverse software ecosystem support at scale. Emerging analytics workloads will not only be able to create user-defined environments suitable to their computational needs independent of batch schedulers or facilities software stacks, but also leverage advanced HPC hardware resources with experiment isolation and management.

I. EXTENDED ABSTRACT

Currently, we are at the forefront of a convergence within scientific computing between High Performance Computing (HPC) and Large Scale Data Analytics (LSDA) [?], [1]. This amalgamation of historically differing viewpoints of Distributed Systems looks to force the combination of performance characteristics of HPC's pursuit towards Exascale with data and programmer oriented concurrency models found in Big Data analytics platforms. Capitalizing upon the community's existing investment in advanced supercomputing systems and economies of scale could benefit both areas beyond what is current possible as disjoint environments. However, current software efforts in each area have become extremely specialized and the gap only continues to grow, making the concurrent support with a single architectural model increasingly intractable.

Instead, we postulate the embracing of software diversity on advanced supercomputing platforms through the use of Virtual Clusters. *TODO: Describe the notion of virtual clusters*

We expect virtualization to be a key aspect to providing Virtual Clusters, however the type and level of virtualization and its interactions with the underlying OS environment are still unknown. Work is needed to determine the most effective way to provide this level of abstraction necessary, and what tradeoffs are necessary in regards to performance considerations, cluster deployment efficiency, OS type and flexibility, workload reproducibility, hardware accessibility, and others. Host virtualization efforts as the Hobbes project [?] provide one example of an OS and virtualization effort that provides the potential building blocks necessary for Virtual Clusters.

OS-level virtualization, or container solutions such as Shifter and Singularity [?], [?] also offer a potential valuable component by extending the notion of Docker by integrating within existing HPC environments. While various virtual abstraction support are important building blocks for Virtual Clusters within advanced technology systems, we still require extensive distributed systems software research and development to overcome the issues brought by batch processing mechanisms. This includes challenges including scheduling & resource management, user-defined image creation & orchestration, network segmentation & isolation, and performance considerations at extreme scale, to name a few.

Currently, workloads dissimilar from traditional parallel HPC simulations, such as emerging data analytics tools and virtualization platforms are forced to work with batch schedulers and shared storage systems that

Virtual Clusters will enable users to focus more on application ecosystem composition matching scientific endeavors rather than forcing the adaption of development environments to platforms that were never designed for such considerations. Effectively, this can lower the barrier of entry to extreme-scale computing for many emerging computational challenges. Furthermore, such virtualized clusters enable more non-standard workflow composition, such as the in-situ coupling of parallel MPI simulations with emerging data analytics and visualization tools for real-time experimental control, either across or within clusters.

Different from clusters in past, not based on commodity systems, but whole new system software architecture is needed.

ACKNOWLEDGMENT

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

REFERENCES

- [1] D. A. Reed and J. Dongarra, "Exascale computing and big data," *Communications of the ACM*, vol. 58, no. 7, pp. 56–68, 2015.