North American Workshop on Silicon Photonics for High Performance Computing

May 17-18 2018, Fort Collins, CO

Managing workload diversity on exascale systems through bandwidth steering

Jeremiah Wilke
Sandia National Labs
jjwilke@sandia.gov

Abstract – Optical transmission cannot be matched for long-distance data movement. However, in supercomputers, high optical bandwidth densities aren't fully leveraged yet. Current supercomputing systems maintained by the US Department of Energy use exclusively electrical packet switches with optical cables only for long-reach links. While torus topologies previously dominated, high-radix packet switches have made hierarchical topologies like dragonfly and fat-tree more popular. Dragonfly topologies, in particular, provide low-diameter connectivity with all-to-all global links between router groups, aiming to minimizing the number of long distance (expensive) optical links. Because of workload diversity (both single applications and multi-job interference), inter-group topologies cannot be designed to fit a single workload or traffic pattern. This means striking a balance between 1) overprovisioning the network to cover all use cases (cost-prohibitive) or 2) underprovisioning the network for certain use cases (performance-prohibitive). Currently, adaptive routing on non-minimal paths through neighboring groups are leveraged to compensate the lack of available bandwidth on direct paths. However, this results in additional network hops, lower power efficiency, and more complicated router implementations. A subtler way to avoid contention involves bandwidth steering with "passive" optical switching. In this talk, we discuss the challenges associated with provisioning networks for the diverse, multi-user workloads on large scientific computing platforms. We then show how optical switches, even low-radix ones, can alleviate some of these challenges through bandwidth steering. We present both past and current work on Flexfly architectures, a dragonfly design incorporating silicon photonics switches that performs such bandwidth steering.

Bio – Jeremiah Wilke is a Principal Member of Technical Staff at Sandia National Laboratories in Livermore, CA, USA. He joined the Scalable Modeling and Analysis group at Sandia in 2011, where he now is one of the lead developers of the Structural Simulation Toolkit (SST). His focus is system-level simulation using the SST macroscale library. His research includes simulation-based co-design of scalable algorithms, distributed memory programming models, and network architectures.