Bandwidth-Aware Resource Management for Extreme Scale Systems

Zhou Zhou[1], Xu Yang[1] and Zhiling Lan[1],
Paul Rich[2], Wei Tang[2], Vitali Morozov[2] and Narayan Desai[23]
[1]Illinois Institute of Technology, [2]Argonne National Laboratory, [3]Ericsson Inc.

Motivations

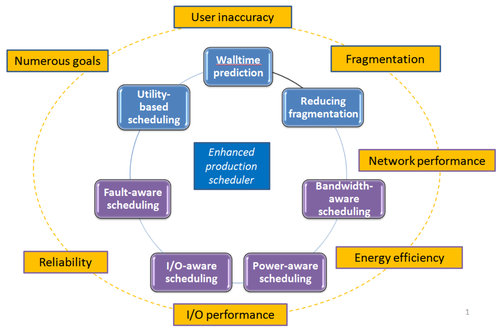
Scheduling challenges on large-scale, production supercomputer:

- Changing and conflicting scheduling goals
- inaccurate job information provided by users (e.g. walltime)
- System fragmentation
- High cost brought by improper scheduling strategy

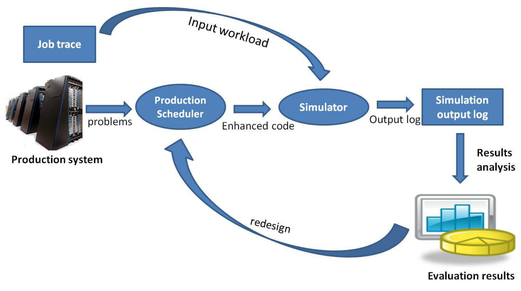
Software challenges on building extreme-scale supercomputers:

- Network performance
- Reliability
- I/O performance
- Energy efficiency

Solution Overview



System Overview



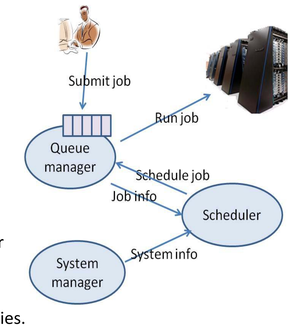
Production Resource Management System

- Cobalt is a open source resource manager developed at Argonne National Laboratory [5].
- Component-based architecture, written in Python

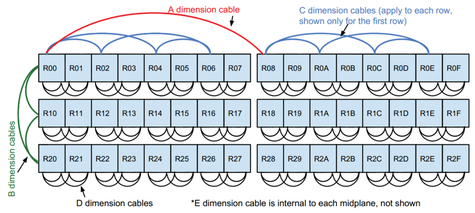
Deployed on a number of Blue Gene systems.

- Support full-system simulation (with Qsim [6])

- QSim is an event-driven simulator for HPC systems [6].
- Trace based simulation
- Support different scheduling policies.



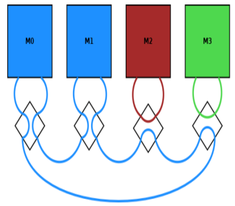
Background of the Blue Gene/Q at Argonne



Flat view of Mira's network topology

Mira is a 10 PFLOPS (peak) Blue Gene/Q system operated by Argonne National Laboratory for the U.S. Department of Energy. Blue Gene/Q is the third generation in the Blue Gene family. Blue Gene/Q systems are comprised of individual racks that are connected together, each rack containing 1;024 sixteen-core nodes, for a total of 16;384 cores per rack. Blue Gene systems have a hierarchical structure.

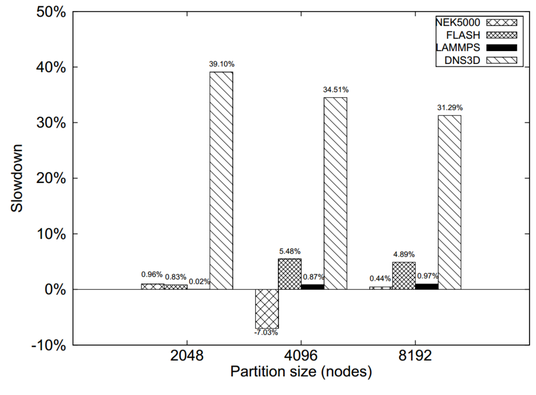
Wire contention between midplanes. A schematic representation of a four midplane long dimension showing a two-midplane torus and two single midplane tori.
Due to the wiring of the 2 midplane torus (blue), and the exclusivity of partitioning on Mira, this prevents the formation of a torus or a mesh with the remaining two midplanes in this dimension. This is representative of both the C and D dimensions on Mira.



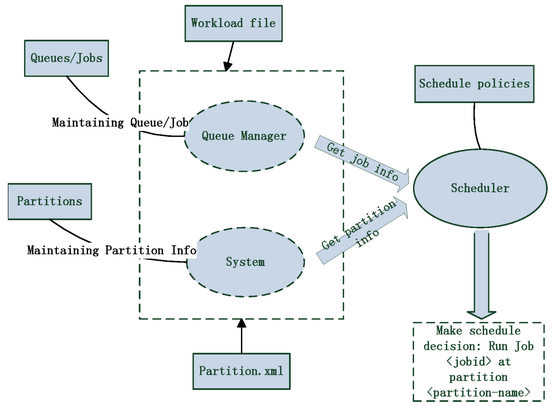
Application Benchmarking

Applications

- NEK5000 is a spectral element CFD code developed at Argonne National Laboratory, which features spectral element multigrid solvers coupled to a highly scalable parallel coarse grid solver [1].
- FLASH 4.0 is the lastest FLASH release from the ASC Center at the University of Chicago. The FLASH code is a multiphysics simulation code written in Fortran90 and C using MPI with OpenMP [2].
- DNS3D is a Direct Numerical Simulation (DNS) code that solves viscous fluid dynamics equations in a periodic rectangular 3-D domain with a pseudo-spectral method of 4th order finite differences and with the standard Runge Kutta fourth order time stepping scheme [3] .
- LAMMPS is a general-purpose molecular dynamics software package for massively parallel computers from Sandia National Laboratories [4].



Simulation Framework



We conduct simulation under a extended open-source event-driven job scheduling simulator called Qsim [6].

- Similar to Cobalt, Qsim contains three components: a queue manager that maintains queued job lists, a system manager that maintains system status, and a job scheduler that makes scheduling.
- Our experimental study is based on real job traces from the Blue Gene/P system and general clusters which represent two popular large-scale systems.
- "Time Stamp" is used to control the progress of the entire simulation including job starting, ending and failure.

Bandwidth-Aware Scheduling Policy

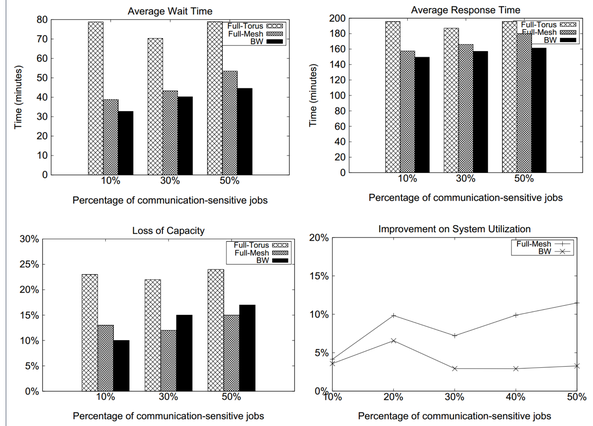
Network Configurations

- Full-Torus: This represents the current network configuration used on Mira where all the partition are fully torus-connected. The partition size ranges from 512 nodes (single midplane) to the whole machine (48 racks).
- Full-Mesh: This configuration is generated from Full-Torus by turning every torus partition into a mesh partition except the 512-node partition that must be a torus. This means wrap-around torus links are turned off in each dimension, which consequently reduce the potential link contention between neighboring partitions.
- Hybrid: This configuration contains a number of mixed partitions: some dimensions are mesh-connected, whereas the others are torus-connected. We call them "minimal contention-free" partitions. On Mira, we build such partitions at 1K, 4K and 32K sizes. We add these mixed partitions into the current full torus network configuration.

Scheduling Policy:

- WFP. WFP favors large and old jobs, adjusting their priorities based on the ratio of wait time to their requested wall clock times. Upon each scheduling iteration, the job at the head of the wait queue is selected to be allocated to a certain partition.
- Least-blocking scheme. LB is used to choose the partition that causes the minimum network contention out of all candidates.
- Bandwidth-Aware Scheduling Policy. Intuitively, jobs running at the problem sizes (such as 1K, 4K and 32K) would be routed to mesh partitions. Jobs exhibiting nontrivial slowdowns should be explicitly submitted to torus partitions of these size. This would allow the majority of jobs to run with smaller resource footprint while still allowing users requiring the full torus partition to have it. The scheme works as follows: in case of communication-sensitive jobs, it will allocate them to torus partitions; otherwise for those communication-insensitive jobs, the scheduler will allocate them to mixed partitions.

Experimental Results



Job Traces

- Job traces used in the simulation are collected from production Blue Gene/Q system named Mir at Argonne National Laboratory.
- The 512-node, 1K and 4K jobs are the majority. For month 2 and 3, 512-node jobs account for half of the jobs.

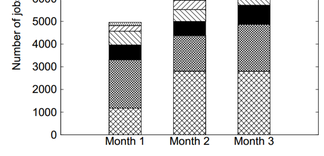


Figure 5. Job size distribution

Reference

[1]P. Fischer, J. Lottes, D. Pointer, and A. Siegel, "Petascale algorithms for reactor hydrodynamics," Journal of Physics: Conference Series, vol. 125, no. 1, p. 012076, 2008.

[2]"The flash code." http://www.flash.uchicago.edu/site/flashcode/

[3]M. Taylor, S. Kurien, and G. Eyink, "Recovering isotropic statistics in turbulence simulations: The kolmogorov 4/5thlaw," Physical Review E, vol. 68, 2003.

[4]"Lammps" https://www.alcf.anl.gov/user-guides/lammps

[5]"Cobalt resource manager." https://trac.mcs.anl.gov/projects/cobalt

[6]W. Tang, Z. Lan, N. Desai, and D. Buettner, "Fault-aware, utility-based job scheduling on blue, gene/p systems," in CLUSTER '09. IEEE International Conference on, Aug 2009, pp. 1–10.

Acknowledgments

This work is supported in part by National Science Foundation grants.

Argonne NATIONAL LABORATORY

NSF