Development of a Computing Cluster at the University of Richmond

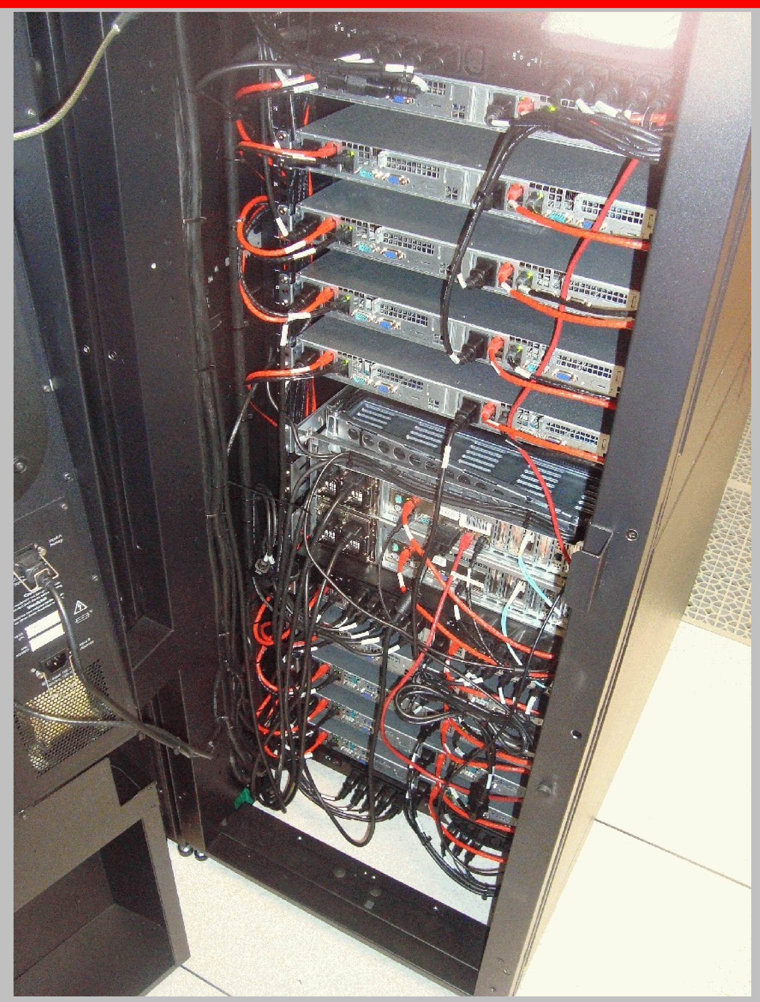
J. K. Carbonneau, G. P. Gilfoyle, and E. F. Bunn

Introduction:

The University of Richmond's new computing cluster was developed to support research programs in both nuclear physics and astrophysics. The nuclear physics research is part of the JLab program, and will provide off-site analysis of data collected at JLab, as well as simulation of the CLAS and CLAS12 particle detectors. The astrophysics research focuses on analysis of cosmic microwave background (CMB) radiation. The cluster will be used by the astrophysics department to analyze data in support of the millimeter-wave bolometric interferometer (MBI) project which is currently funded by NASA.

Overview:

The cluster is a $120,000 system consisting of 20 nodes, each with 12 Westmere cores. It is currently running Red Hat Enterprise Linux 5. In order to distribute a single job across all nodes, the cluster uses a Message Passing Interface (MPI). The MPI keeps track of the status of the individual cores on each node, and handles the distribution of a submitted job across nodes. In order to submit jobs, we used a script called QSUB, which submits a job to the cluster, reserving for it's use a certain number of cores and nodes. This allows many jobs to be run at once, each one only using the resources it needs, and otherwise waiting in a queue for resources to become available if they currently are not.



Each slot on the rack holds two of the cluster's remote nodes. Each remote node has dual six-core X5650 Westmere 2.66 GHz processors, 24GB of RAM, and 2 500GB hard drives in a RAID-0 array. The head node has 24GB of RAM and 4 500GB drives in a RAID5 array, and the storage node has 12GB of RAM with 8 1TB drives in RAID5.

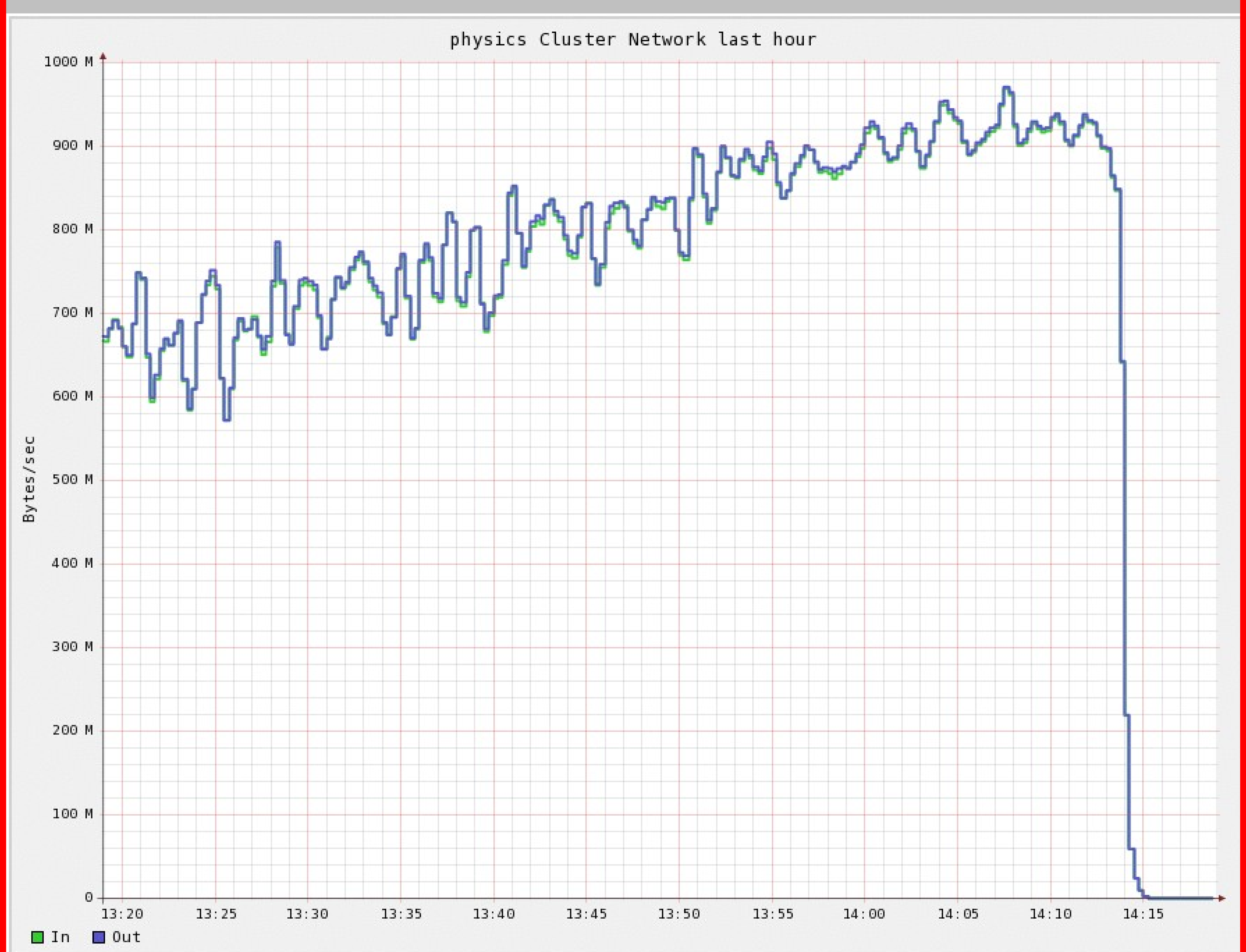


Fig 1. This is a graph of the cluster network activity in bytes per second over one hour during a benchmark test. The traffic gradually approaches the maximum network speed (1000 MB/sec), creating a performance bottleneck.

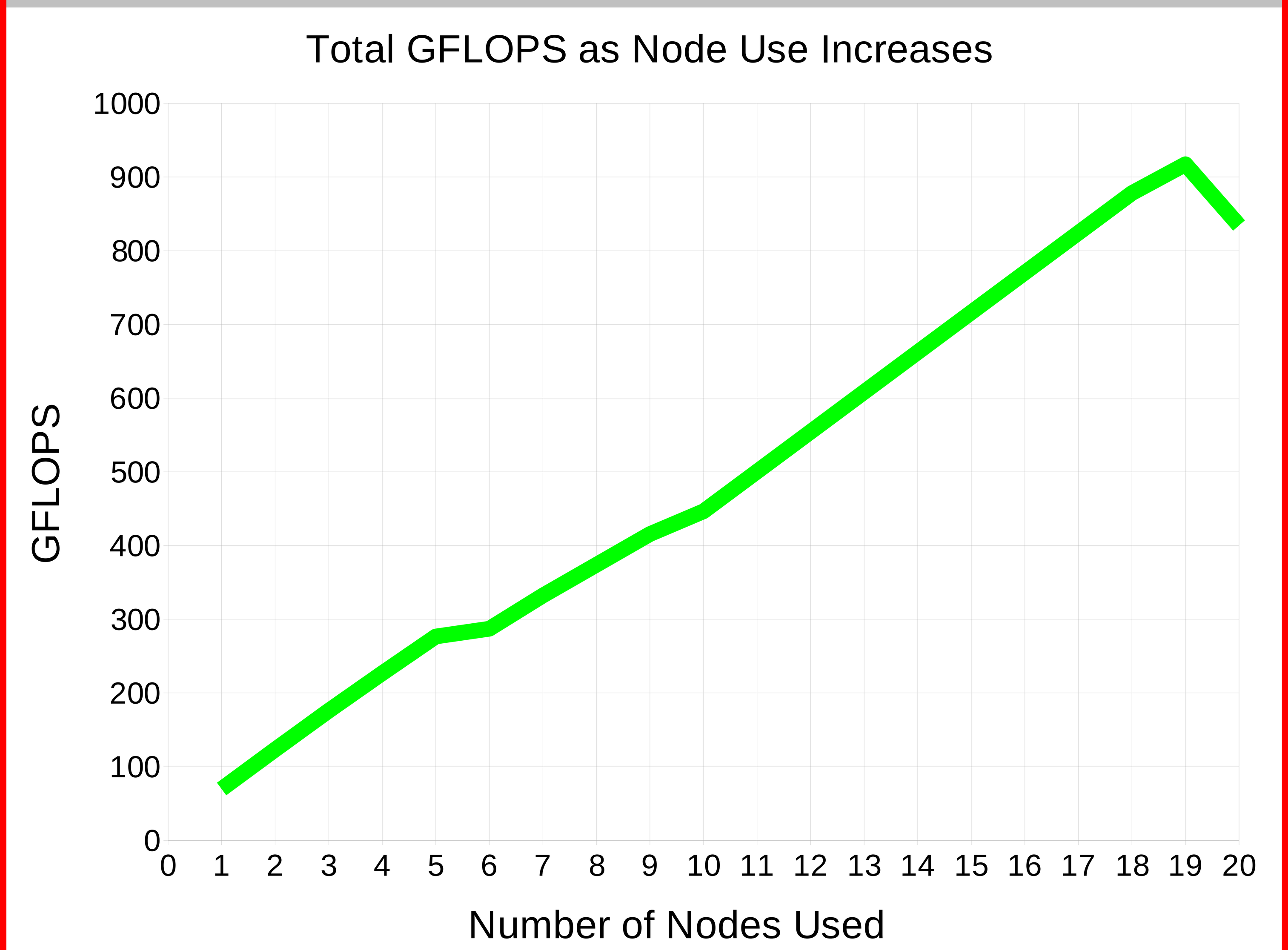


Fig 2. This graph shows the results of a benchmark test that used an increasing number of nodes for each test. The system peaked at 917.3 GFLOPS when using 19 of the 20 nodes. The drop for 20 nodes is because when all nodes are used at once, performance suffers from a bottleneck in communication speed between the nodes.

Efficiency:

In order to test the performance of the cluster, we used the High Performance Linpack (HPL) benchmarking program, which is a standard for benchmarking high-performance machines. Based on the number of nodes and the clock speed, the theoretical peak performance of the system is about 2500 GFLOPS. However, testing with HPL showed a maximum speed of 917 GFLOPS. This discrepancy is due to the network traffic between the nodes, which approaches it's maximum speed as MPI passes large amounts of data between the nodes, becoming a bottleneck (see Fig 1). The current top super computer in the world, the Cray Jaguar, is a $104 million system with a peak performance of 2331000 GFLOPS, giving it a cost efficiency of $44 per GFLOP. Our system cost about $48 per GFLOP, making it only $4 more per GFLOP than the worlds fastest computer.

Applications:

The cluster is currently being used by the nuclear physics and cosmology departments at the University of Richmond, and is running several programs to this end.

GEMC

gemc (GEant4 MonteCarlo) is a C++, object-oriented simulation of the future CLAS12 detector built on the Geant4 package from CERN. It reads detector information from an external database (mysql), hits in detector components are built by a factory method at run time, and the output banks are dynamic. Qt4 is used for a graphical user interface.

Pythia

The Pythia program is a simulation of high-energy collisions built on a selection of physics models for the evolution from a few-body hard process to a complex multihadronic final states. It can be used to calculate initial- and final-state parton showers, multiple parton-parton interactions, beam remnants, string fragmentation and particle decays.

ROOT

ROOT is an C++-based, object-oriented program and library developed by CERN designed for particle physics data analysis. It provides tools for (among other things) histogramming, graphing, curve fitting, statistics, and four-vector computations. It is routinely used as the final step in the analysis chain for gemc, PYTHIA and other Monte Carlo simulations.

MADmap

MADmap generates a likelihood map of the sky from time ordered data from an instrument. It is written to be run on parallel systems, making it ideal for being run on the cluster.