

Real Application Performance with the InfiniPath InfiniBand Interconnect

Real Application Performance with the QLogic InfiniPath InfiniBand Interconnect

By Greg Lindahl, Chief Scientist, QLogic System Interconnect Group

Executive Summary

The best benchmark of a supercomputer (or any computer) is always your own application. In the real world, users always end up relying on standard benchmarks to figure out which systems are interesting enough to be worth benchmarking with the user's own application. Individual processors such as the AMD Opteron and Pentium4 are intensively benchmarked on serial applications, and standard benchmarks such as the SPEC cpu benchmarks are easily available and useful for a "first cut" at application performance. Parallel benchmarks, especially of MPI applications, have proven to be much more difficult to find. It is fairly easy to find the "hero numbers" – micro-benchmarks of latency and bandwidth – for an interconnect. But these "hero numbers" are often misleading, and do not accurately predict real application performance.

This whitepaper contains real application performance data on clusters scaling out to 512 cpus for QLogic's InfiniPath InfiniBand interconnect, compared with data from competing interconnects. In most cases, the benchmark code, test case, and competing data are publicly available. The applications are representative of code used in many scientific and engineering disciplines. The test cases are mostly chosen by the application authors as being representative of actual application usage.

The results are quite striking. In nearly all cases, the InfiniPath cluster outperforms clusters using competing interconnects, such as products from Myrinet, Mellanox, Quadrics, SGI, and IBM. In all cases, the price/performance of the InfiniPath InfiniBand cluster is much better than competing solutions. Finally, we present several dual-core Opteron results, which show that dual core offers good performance and excellent price/performance.

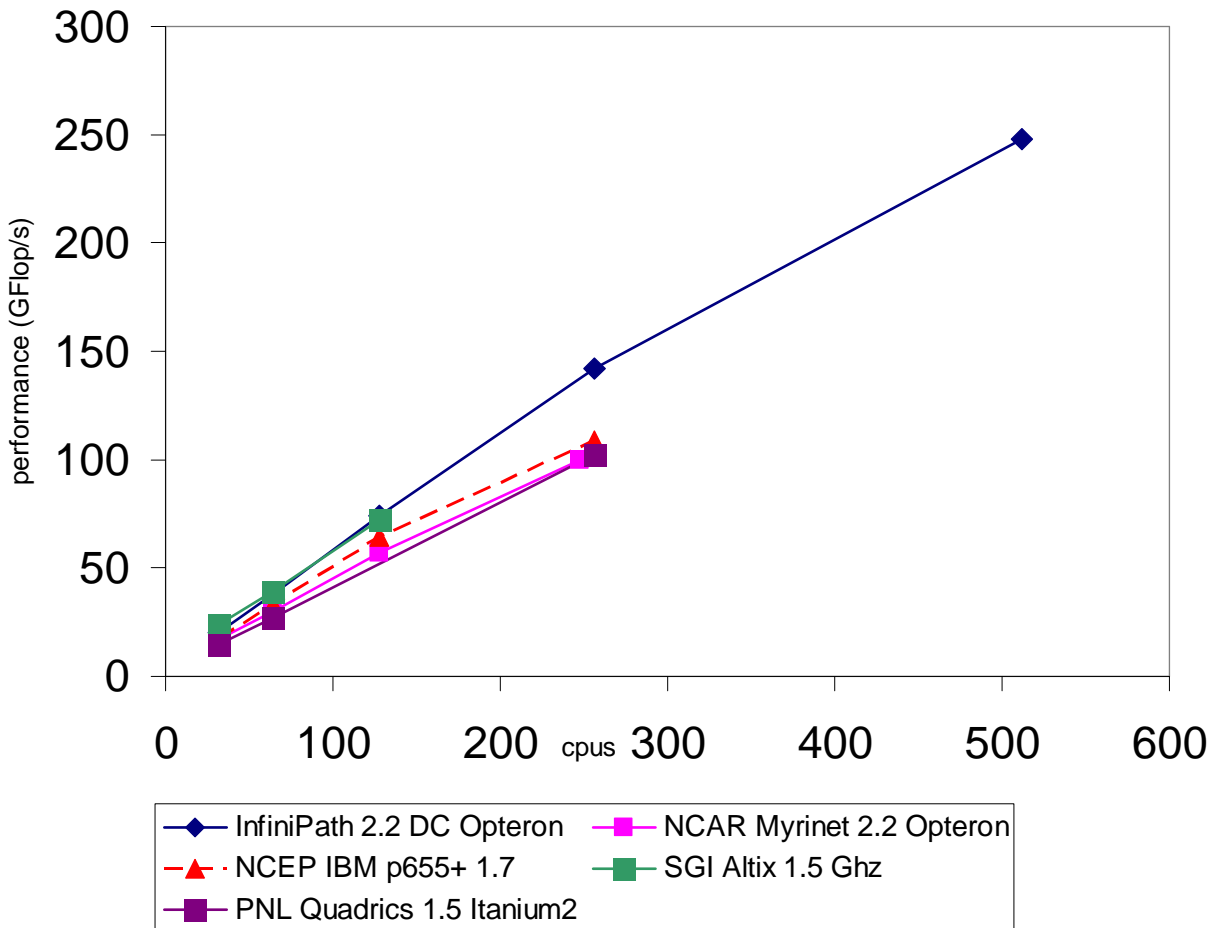
We hope that the availability of this data leads to more data being available. If you run these benchmarks on your current clusters, we'd like to hear the results. If you'd like to verify any of our results for yourself, or test your own application on an InfiniPath cluster, please contact sales@QLogic.com. We have clusters ranging from 32 cpus to the 576-core "Emerald" cluster at AMD's Developer Center available for customer benchmarking.

2 Real Application Performance with the QLogic InfiniPath IB Interconnect

WRF 2.0

WRF (pronounced like “wharf”) is a next-generation weather code used for both mesoscale and continental US weather prediction. It is a joint project between weather forecasting and research organizations including NCAR, NCEP, FSL, AFWA, NRL, Oklahoma University, and the FAA. WRF is currently used for operational weather forecasts at the US National Weather Service.

The results shown below are for a highly-documented test-case named 2001102400. Competing data was run by a variety of organizations involved in the WRF project, and can be found at <http://www.mmm.ucar.edu/wrf/bench/wrf-bench.bmp>.



The graphic shows that the InfiniPath-equipped 2.2 Ghz dual-core Opterons (a cluster at AMD’s Benchmark Center named “Emerald”) starts off slightly ahead at small cpu counts.

3 | Real Application Performance with the QLogic InfiniPath IB Interconnect

At 256 cpus, the InfiniPath cluster has pulled away from the pack and provides unprecedented scaling on WRF 2.0. Note that these are dual-core AMD Opterons; the price of the InfiniPath Opteron cluster is much less than most of the competing solutions presented here. QLogic will run the WRF 2.1 benchmark as soon as it is finalized.

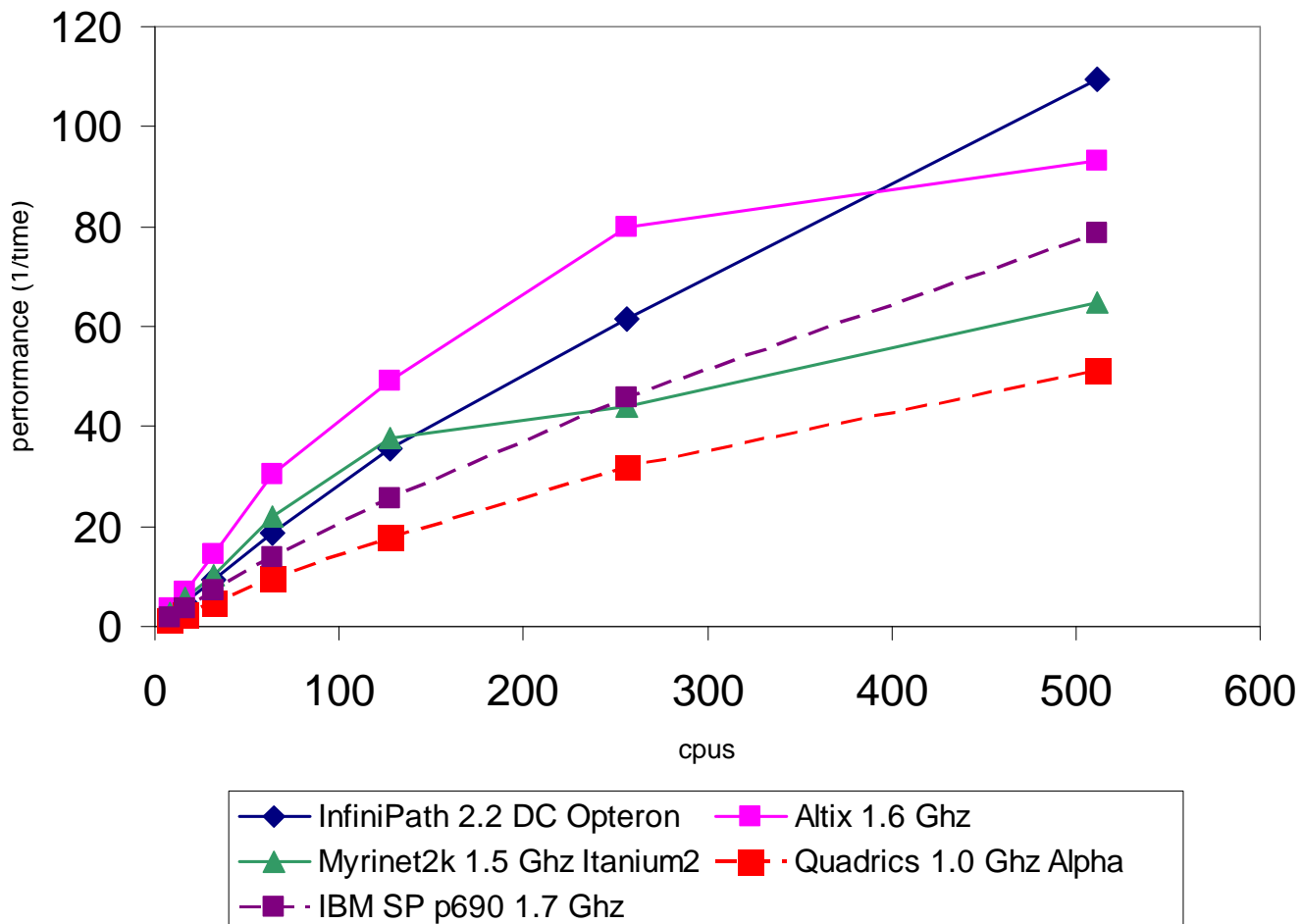
4 Real Application Performance with the QLogic InfiniPath IB Interconnect

NAMD

NAMD is a widely-used molecular dynamics code.

This graph uses data from <http://www.ks.uiuc.edu/Research/namd/performance.html>

As you can see, the scalar advantage of the very expensive SGI Itanium2 Altix system is eventually outperformed by the superior scaling of the InfiniPath InfiniBand interconnect. A graph of performance verses price would show the InfiniPath system ahead at all cpu sizes.

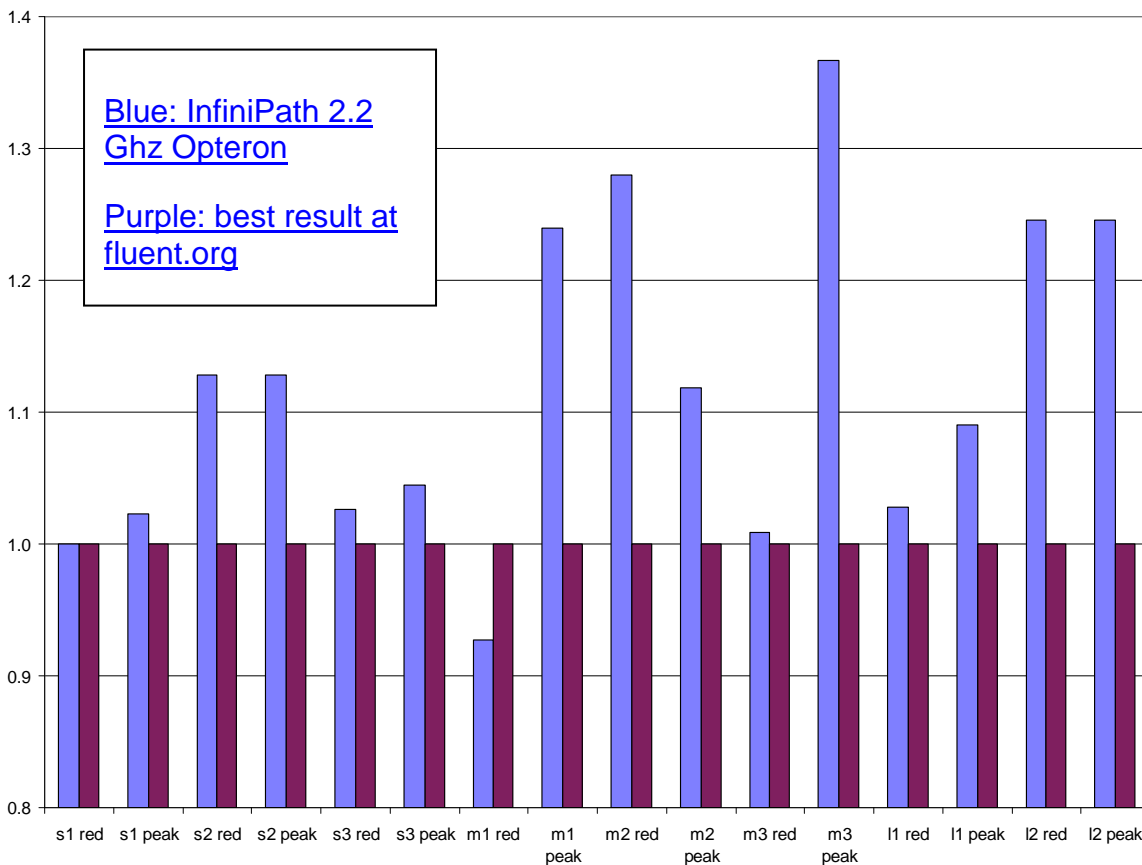


5 Real Application Performance with the QLogic InfiniPath IB Interconnect

Fluent

Fluent is a leading commercial application used for CFD (computational fluid dynamics). Fluent's website contains 9 benchmarks with data of small, medium, and large sizes, which represent the types of computations done by Fluent's users. Results are presented as a 2 "ratings": one is the maximum performance which can be attained at 67% or better parallel efficiency, called the "red zone", and the other is the peak performance without regard to parallel efficiency.

The best result collected at <http://fluent.org/software/fluent/fl5bench/> was beaten almost all of the time by the InfiniPath-equipped cluster. The results which were beaten came from a variety of systems, including Opteron systems with Mellanox interconnect, EM64T systems with Mellanox interconnect, and a 1.5 Ghz SGI Altix.



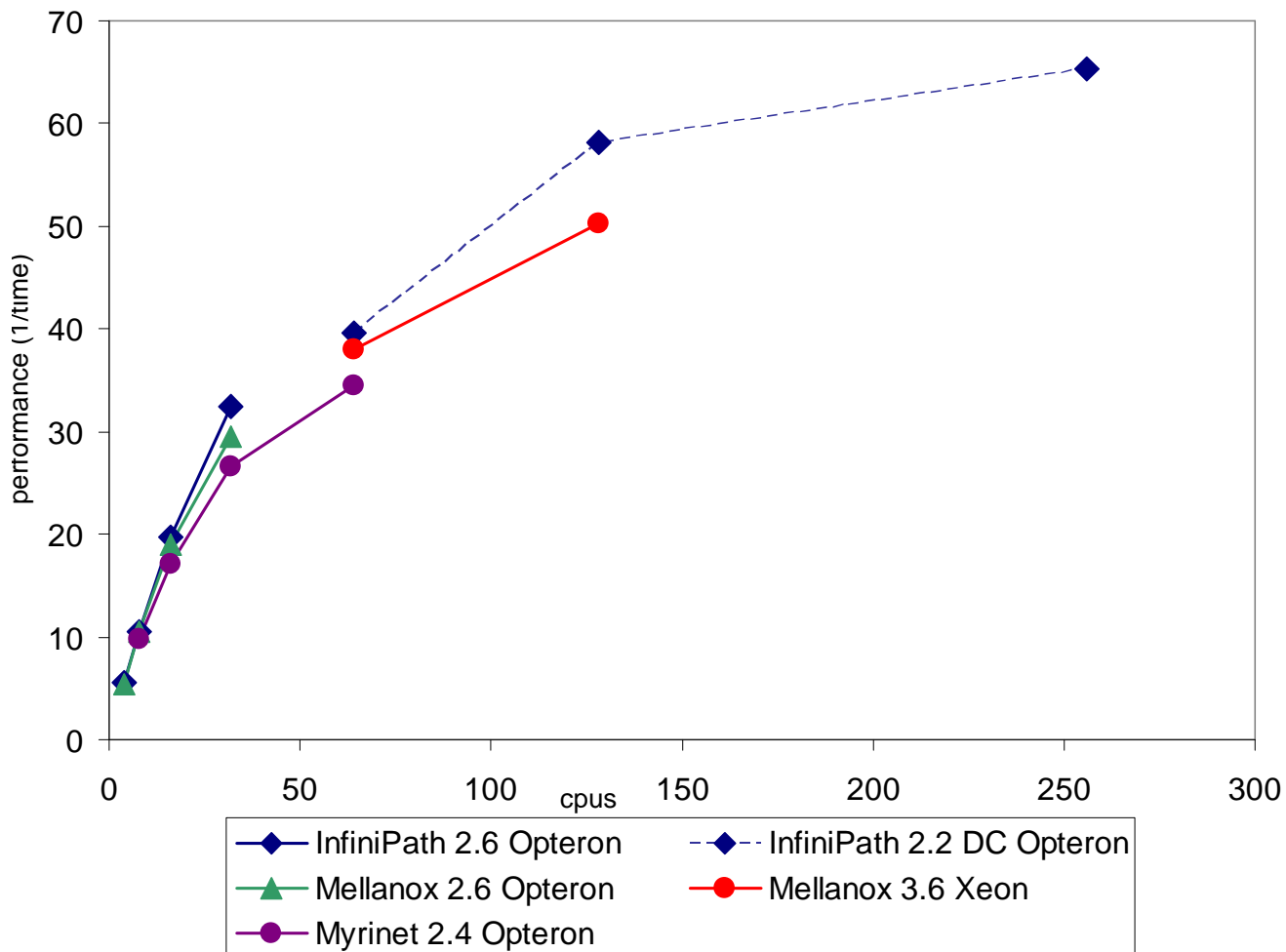
Note: A commercially supported version of Fluent for InfiniPath is not currently released at the time of publication.

6 Real Application Performance with the QLogic InfiniPath IB Interconnect

LS-Dyna

LS-Dyna from LSTC is a commercial Finite Element Analysis (FEA) application, frequently used for automotive crash-testing simulations, and other finite-element uses. The website <http://topcrunch.org/> contains benchmark results for several problems. The graphic below shows results for the “neon refined” test case. All data graphed is available at the “topcrunch” website.

The graph shows the InfiniPath-equipped cluster winning at both small and large sizes, verses both Myrinet and Mellanox-based Opteron & Xeon clusters. Due to these results, LSTC has recently announced formal support of the InfiniPath InfiniBand interconnect.

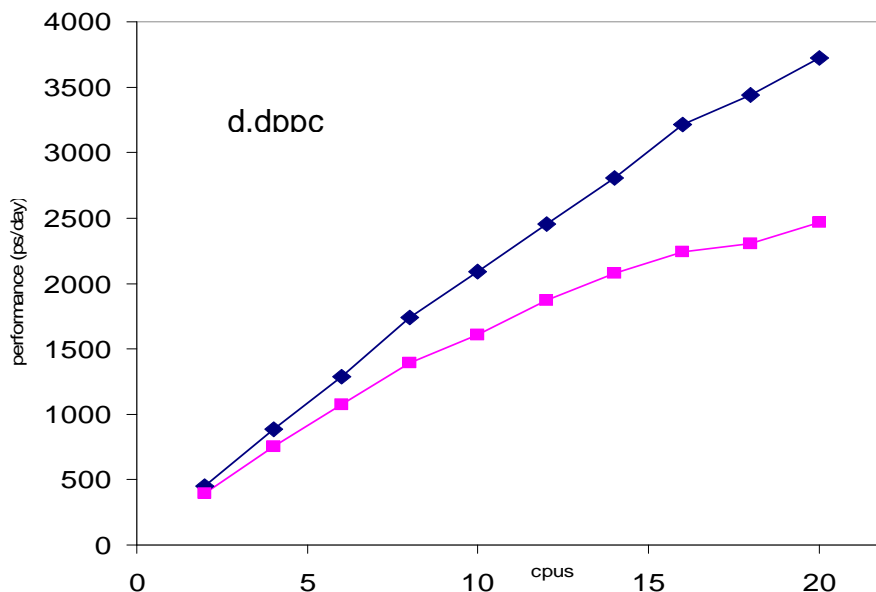
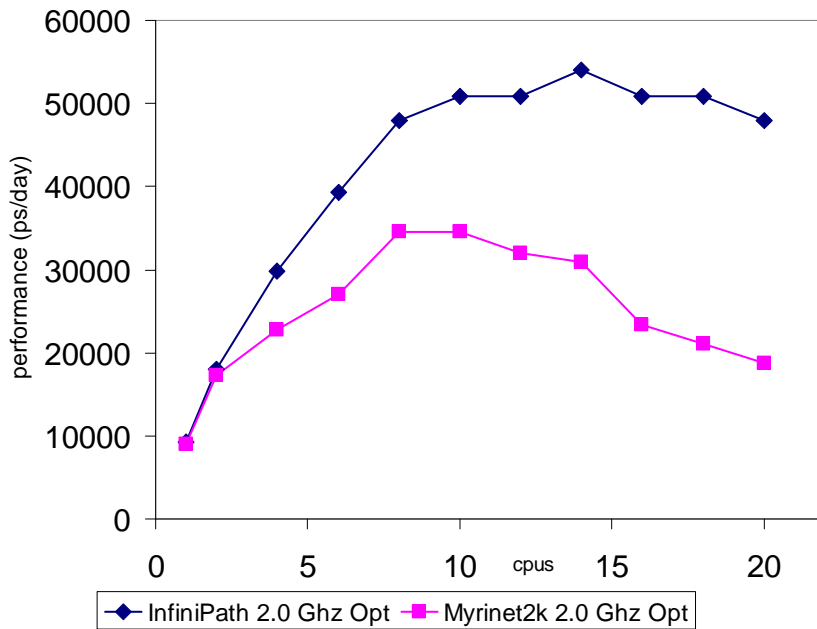


7 Real Application Performance with the QLogic InfiniPath IB Interconnect

Gromacs

Gromacs is a commonly-used molecular dynamics code. The two test cases below are both difficult to scale. In “d.villin,” the both the Myrinet and InfiniPath-equipped clusters abruptly stop scaling around 8 cpus. However, the performance achieved at this cpu count is approximately 40% higher with InfiniPath than a Myrinet2k-equipped cluster with similar 2.0 Ghz Opteron cpus. In the other test, “d.dppc,” the Myrinet2k-equipped cluster is reaching its scalability limit around 20 cpus, while the InfiniPath-equipped cluster is both higher performance and is still headed up. See <http://biowulf.nih.gov/apps/gromacs/bench.html>

Gromacs performance is measured in “ps/day”, pico-seconds of simulation time per day of compute time.



8 Real Application Performance with the QLogic InfiniPath IB Interconnect

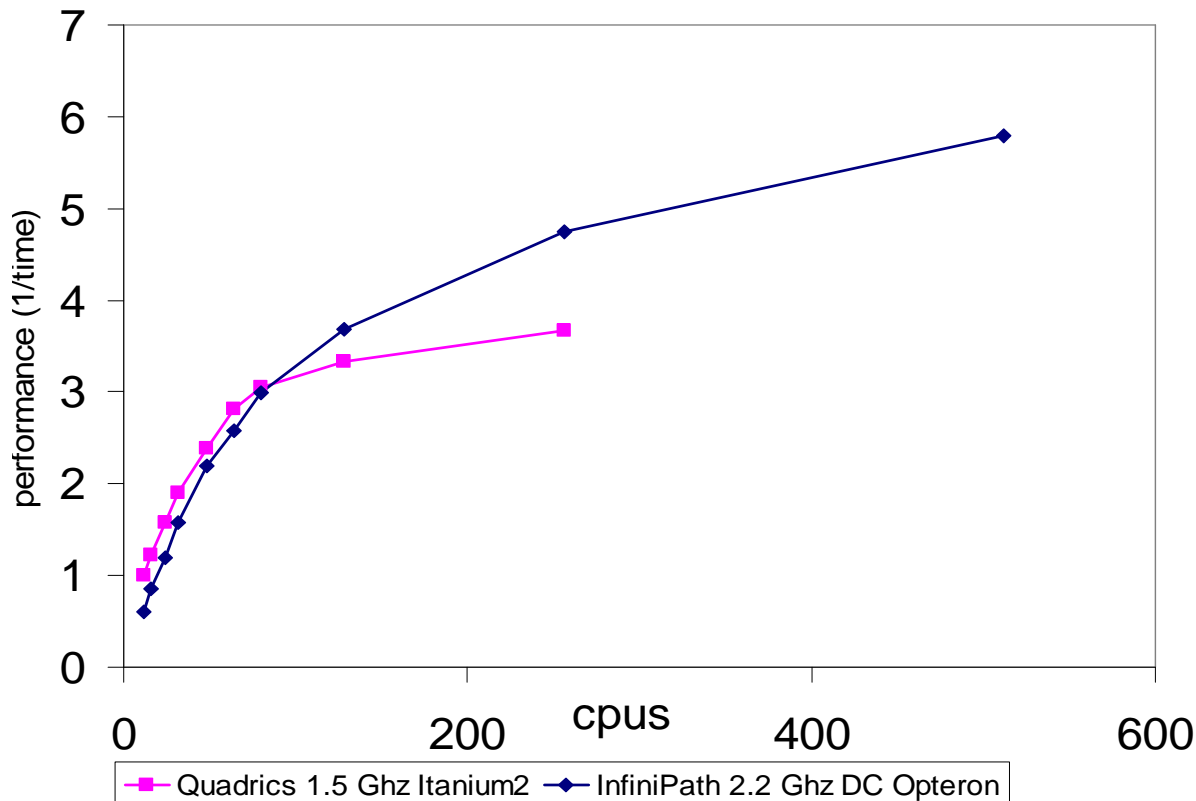
NWChem

NWChem is a computational chemistry code developed at the Molecular Sciences Software group of the Environmental Molecular Sciences Laboratory at the Pacific Northwest National Laboratory. NWChem is built on top of a one-sided messaging library named ARMCI, and many interconnects such as Quadrics have specially-written ARMCI drivers.

We modified ARMCI to implement one-sided operations using two-sided MPI messages. While this isn't the most efficient implementation, it was good enough to allow an InfiniPath-equipped cluster to out-scale PNNL's Itanium2 + Quadrics cluster. PNL's data is in pink, the InfiniPath 2.2 Ghz dual-core Opteron cluster (AMD Emerald) is in blue. As you can see, the Itanium cluster starts out at a substantial performance advantage, perhaps due to NWChem's heavy use of matrix algebra. The superior scalability of the InfiniPath interconnect pulls into the lead above 64 cpus. Note that these are dual-core Opterons; the price of the InfiniPath Opteron cluster is less than one-half the price of this Itanium/Quadrics cluster.

This test case is Siosi6. Comparable results may be found at:

<http://www.emsl.pnl.gov/docs/nwchem/benchmarks/>



NWChem Citation

NWChem's license requires the following citation in any paper giving NWChem results:

Aprà, E.; Windus, T.L.; Straatsma, T.P.; Bylaska, E.J.; de Jong, W.; Hirata, S.; Valiev, M.; Hackler, M.; Pollack, L.; Kowalski, K.; Harrison, R.; Dupuis, M.; Smith, D.M.A.; Nieplocha, J.; Tipparaju V.; Krishnan, M.; Auer, A.A.; Brown, E.; Cisneros, G.; Fann, G.; Fruchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J.; Tsemekhman, K.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z.; "NWChem, A Computational Chemistry Package for Parallel Computers, Version 4.7" (2005), Pacific Northwest National Laboratory, Richland, Washington 99352-0999, USA. A modified version.

"High Performance Computational Chemistry: an Overview of NWChem a Distributed Parallel Application", Kendall, R.A.; Aprà, E.; Bernholdt, D.E.; Bylaska, E.J.; Dupuis, M.; Fann, G.I.; Harrison, R.J.; Ju, J.; Nichols, J.A.; Nieplocha, J.; Straatsma, T.P.; Windus, T.L.; Wong, A.T. Computer Phys. Comm., 2000, 128, 260-283

For More Information

Additional information on QLogic InfiniBand products can be obtained by visiting <http://www.qlogic.com/pathscale> on the World Wide Web, or by contacting:



QLogic, inc.
System Interconnect Group Phone: (650)-934-8100
2071 Stierlin Court, Suite 200 Fax: (650) 428-1969
Mountain View, CA. 94043 USA www.qlogic.com/pathscale

Copyright 2005 PathScale, Incorporated.

QLogic, PathScale and InfiniPath are trademarks of QLogic, Incorporated.
AMD and AMD Opteron are trademarks of Advanced Micro Devices, Incorporated.
All other trademarks and registered trademarks are the property of their respective owners.