

HP Clusters in Computational Chemistry

Case Study with Amsterdam Density Functional (ADF) Software



Executive Summary	2
Introduction.....	3
DFT Calculation for a Metallophthalocyanine: A Typical Example	3
Comparative Results	6
Parallelism	6
Computing Architectures	7
I/O and File System Access	8
Interconnects	9
Conclusion: Reducing Overall Computational Time	10
Glossary of HP Terms	11
For more information	12

Executive Summary

Effective use of cluster systems may simultaneously speed computational chemistry research and offer cost savings. Many factors may contribute to its effective implementation. Using a realistic large molecule as a case study, HP and Scientific Computing & Modeling (SCM) characterized the cluster configuration requirements of a Density Functional Theory (DFT) calculation with ADF, SCM's quantum chemistry program.

SCM supports HP-MPI for its optimized cluster implementation of ADF.

HP-MPI provides transparent application support of different clusters allowing developers and users to run the same MPI-based applications across multiple platforms and interconnects without the headaches of recompilation.

This study reveals excellent scaling, independent of architecture, is found for HP clusters (tested up to 128 processors), effectively reducing the overall computational time. For DFT calculations, clusters of Intel® Itanium® 2 servers, such as HP Cluster Platform 6000, have a distinct performance advantage over similar clusters based on Intel® Xeon® processors, which in turn outperform clusters based on AMD® Opteron™ processors.

Cluster performance, even with the fastest computing architecture, is dependent on sufficient memory and I/O. ADF is a memory-intensive application and, for some job types, stores large amounts of data on disk. Using today's typical configurations, this case study reveals memory was not a bottleneck to performance. I/O, on the other hand, was improved (from typical Network File System (NFS) configurations) with HP StorageWorks Scalable File Share (SFS)—delivering the best I/O performance, especially useful in large cluster configurations.

Performance on clusters may also be limited by interconnect speed. In this case study, good parallel speedups of clusters with more than four processors came from using high performance interconnects.

The parallel efficiency of a particular ADF job depends on many factors. As a guide, this case study characterized the effect of a number of key cluster configuration variables on cluster performance. It also represents part of the collective analyses and understanding that HP and its partners develop in better supporting its computing customers in life and materials sciences.

Introduction

Typically, computational chemistry research projects require the running of numerous computational jobs, some of which may run for days, weeks, or months. Clusters may offer substantial time and budget savings if the software takes advantage of the distributed parallelism of cluster technology. Thus, understanding the factors affecting cluster performance becomes critical for researchers pushing the limits of their computational resources to study larger and more complex molecules. To that end, this paper describes a joint investigation of the parallel performance of SCM's Amsterdam Density Functional (ADF) software on a number of configurations of HP Linux clusters. ADF is a widely used quantum chemistry program known for its efficient density functional methods.

SCM has built into ADF software a number of features to reduce overall computational time, including the use of full molecular symmetry, distance cutoffs to achieve linear scaling with system size in performance critical parts of the code, the option to use frozen-core basis sets, and the use of an atom-pair based density fit. It has also been written to take advantage of compiler options and computer hardware features, and has been optimized to run on HP Linux clusters. In addition, users may further improve the performance of ADF on their clusters by reducing or eliminating bottlenecks on their systems.

Clusters are excellent for substantially reducing total job completion times for software that has been parallelized. A typical ADF job on a single processor spends most of its time doing numerical integrations. These integrations involve a grid that contains thousands of integration points per atom. The straightforward and effective way to parallelize ADF is to subdivide this numerical grid, so that available processors perform integration with minimal need for communication.

The density fit takes most of the remainder of the time spent in a DFT calculation using ADF; this is done in a loop over atom pairs. These calculations over pairs are readily distributed over multiple processors to achieve efficient scaling.

Parallel efficiency is currently limited by diagonalization, which is performed in serial mode in ADF. There are plans to overcome this limitation with the use of the MPI parallel ScaLAPACK library.

Parallel performance may also be limited by load imbalances. In the density fit, some atom pairs (such as heavy atoms with many fit functions) have a higher computational load than others, so the processes are given uneven loads. Parallel performance is determined by the slowest process, which is typically the one that has to do the most work. Slowdown may also occur if one part of the integration grid requires more computations, due to its central position in the molecule, than the rest of the integration grid. ADF is written to take these potential imbalances into account, but currently cannot fully eliminate them.

The ADF software program

ADF focuses on electronic structure calculations based on Density Functional Theory (DFT). The DFT approach has earned a dominant position within the quantum chemistry community. This approach combines efficiency (systems with hundreds of atoms can be treated) with accuracy, and applies well to molecules that are difficult to handle with Hartree-Fock based methods.

The ADF code consists of a periodic structure code BAND¹ to study systems such as polymers, solids, slabs, and surfaces. ADF is especially useful for the treatment of finite systems, such as transition metal compounds, molecules containing heavy elements, metal clusters, proteins, and molecules in a solvent environment.

In addition to standard calculations such as geometry optimization, transition state search, and analytical frequency calculations, ADF, which uses the Slater type basis sets, can calculate a wide variety of molecular properties, such as optical spectra, NMR¹, ESR¹, and hyperfine interactions. It can also deal with many types of environments, using QM/MM¹ for protein environments, or COSMO¹ and discrete models to account for solvent effects.

ADF is developed by a number of academic research groups in collaboration with SCM. Most developments in ADF have come from the Baerends group in Amsterdam, The Netherlands, and the Ziegler group in Calgary, Canada. For more information, see <http://www.scm.com>.

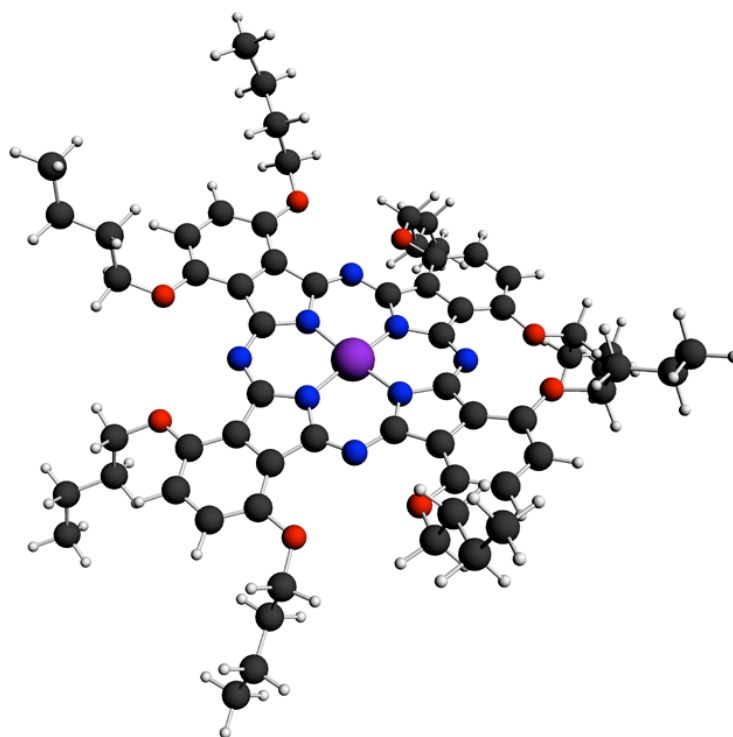
DFT Calculation for a Metallophthalocyanine: A Typical Example

Metallophthalocyanines have been studied extensively since the beginning of the 20th century because of their interesting properties: they have high extinction coefficients so thin films of these molecules offer bright colors; they are both photo and thermally stable; they offer choices of a number of different absorption bands and change color relatively quickly. These properties make metallophthalocyanines potential materials for applications in the fields of molecular electronics, liquid crystals, and photovoltaic solar cells. Most recently, these compounds have gained attention as photothermal sensitizers for tumor therapy purposes.

A metallophthalocyanine, NiPc(BuO)₈ [Octabutoxy Phthalocyaninato-Nickel (II)] was chosen for this case study (Figure 1). Understanding the factors governing the efficiency of NiPc(BuO)₈'s photothermal behavior requires a knowledge of its excited state dynamics.

To elucidate this, Time-dependent Density Functional Theory (TDDFT) is used to calculate excitation energies and oscillator.

Figure 1. The study molecule: NiPc(BuO)₈



The results were compared with those from an earlier study by Gunaratne et al. (J. Phys. Chem. A, Vol. 109, No. 10, 2005 p. 2078) who used ADF to study a model compound, NiPc(MeO)₈ where the butyl groups were replaced by methyl groups in order to make their calculations more tractable. The model's experimental results are consistent with the results found with this case study's calculations, with minor changes in the spectra (shifts below 0.1 eV) due to the replacement of the butyl groups by the methyl groups. Both sets of results also agree with experimental data.

As a case study, NiPc(BuO)₈ represents a realistic large example, one that does not atypically favor scaling (normal input settings were used). It has 161 atoms and D_{2d} symmetry. All 30 lowest excitations were calculated using Becke-Perdew/Adiabatic Local Density Approximation functionals with a large triple-zeta basis set with two sets of polarization functions (TZ2P). The TDDFT calculation typically takes about one day of computation on a single state-of-the-art processor.

To study the characteristics of ADF's TDDFT module, this calculation was run in a number of cluster computing environments, varying the number of processors, the processor architecture, I/O and interconnect.

To improve the ease of running ADF on clusters, SCM adopted HP-MPI. Unlike MPICH and other common flavors of MPI, HP-MPI enables ADF to run on clusters with different interconnects without the need to recompile. The use of HP-MPI by ADF eliminates the headaches associated with recompilation due to link problems and issues resulting from differing compiler versions. However, since communications are not a significant part of these ADF runs, a higher performance interconnect has little effect.

Table 1. The calculation was performed on the following HP Cluster Platform (CP) configurations. For a complete description of the product terminology, please see the [Glossary](#).

HP CP 6000 Architecture: Intel Itanium 2, 1.5 GHz Node: 2 processor 2 core rx2600 server Operating System: (Linux) XC v. 3.0 Interconnect: ELAN 4
HP CP 6000 Architecture: Intel Itanium 2, 1.6 GHz Node: 2 processor 2 core rx1620 server Operating System: (Linux) XC v. 3.0 Interconnect: InfiniBand SDR or Gigabit Ethernet
HP CP 4000 Architecture: AMD Opteron, 2.2 GHz Node: 2 processor 2 core HP ProLiant DL145 server Operating System: (Linux) XC v3.0 Interconnect: InfiniBand SDR
HP CP 3000 Architecture: Intel Xeon, 3.4 GHz Node: 2 processor 2 core HP ProLiant DL360 G4 server Operating System: (Linux) XC v. 3.0 Interconnect: Myrinet D

All the cluster configurations are part of the HP Unified Cluster Portfolio (UCP) that combines the flexibility of a custom solution with a preconfigured, pre-tested, factory-built product. HP's UCP and its Cluster Platforms (CP) are specifically targeted at high performance computing users to allow them to focus on running their jobs on clusters of 5 to 1024 nodes.

XC System Software, supported by HP, is based on a standard Linux (Red Hat Enterprise Linux 4.0 Update 2) distribution combined with both commercial and open source packages. Included in this is support for both Network File System 3 (NFS) and HP StorageWorks Scalable File Share (SFS), a powerful cluster file system that gives users of Linux clusters scalable storage (scaling to over 35 GB/sec aggregate bandwidth with a capacity that scales from 1 TB to 512 TB) that is easy to use and easy to administer.

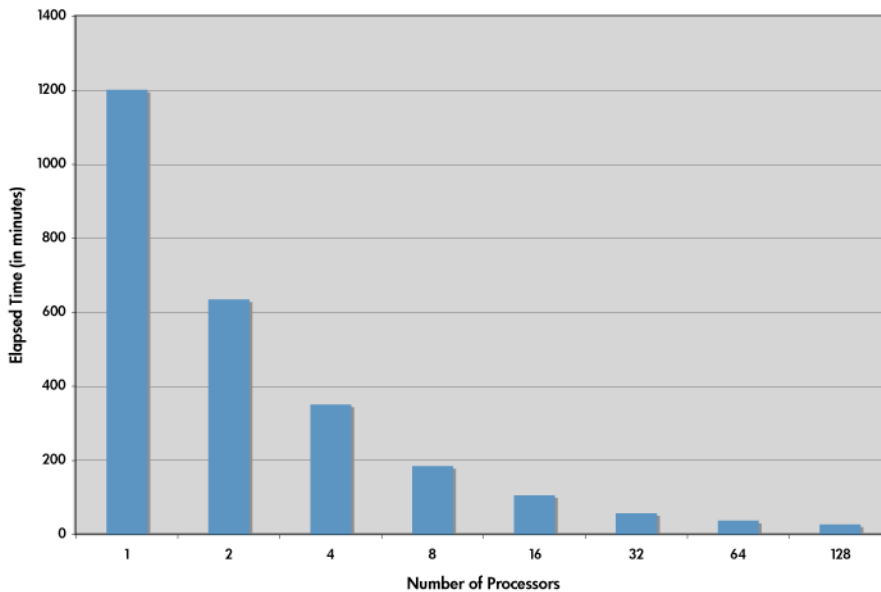
For more information on HP Unified Cluster Portfolio and the individual cluster solutions, go to <http://h20311.www2.hp.com/HPC/cache/275420-0-0-121.html>

Comparative Results

Parallelism

The efficiency of scaling is measured by running the test calculation on various HP Linux clusters with up to 128 processors. The results reveal excellent scaling (graphs 1 and 2) for all three architectures tested. No significant differences in scaling appeared among the three architectures tested (graph 2). For up to 32 processors, speed-up in performance is almost perfect – in the range of 1.8 to 2.0 when doubling the processor count. From 32 processors to 64 processors, scaling falls off to between 1.5 and 1.6. Going from 64 to 128 processors, scaling falls off further to 1.4.

Graph 1. Scaling of HP CP 6000 cluster of HP Integrity rx2600 servers from one to 128 processors.



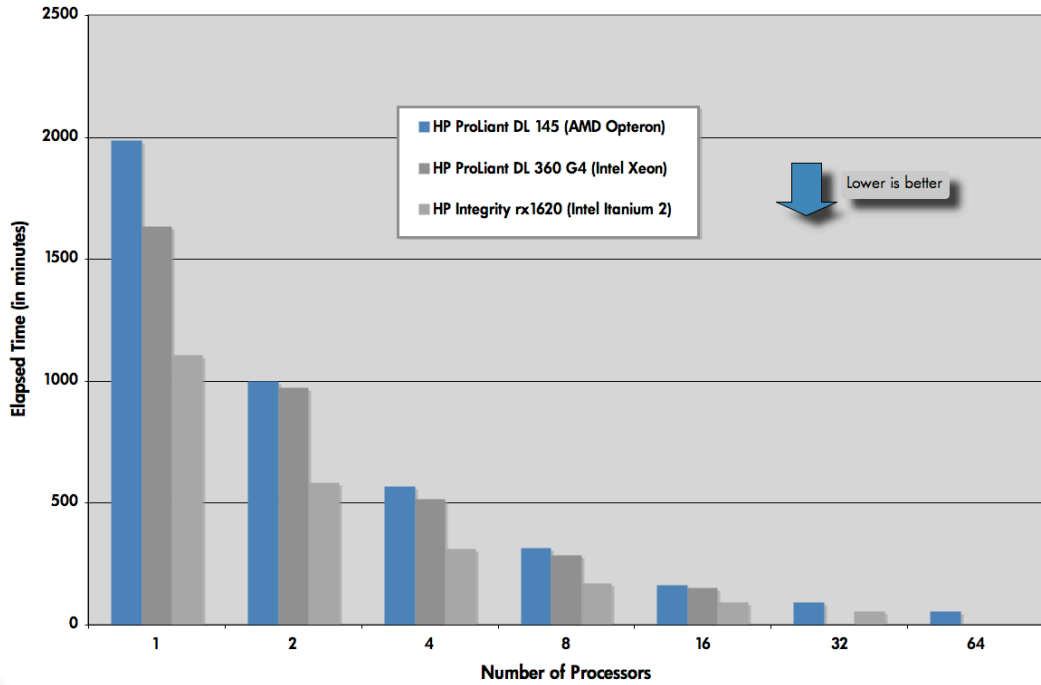
Configuration details for Graph 1

HP CP 6000
Architecture: Intel Itanium 2, 1.5 GHz
Node: 2 processor 2 core rx2600 server
Operating System: (Linux) XC v. 3.0
Interconnect: ELAN 4
Disk: HP SFS

Computing Architectures

Graph 2 also shows the relative performance of three computing architectures: Intel’s Itanium 2 and Xeon DP processors and AMD’s Opteron 200 Series processor. The Itanium 2 based (1.6 GHz) cluster is about 1.6 times faster than the Intel Xeon (3.4 GHz) based cluster and 1.8 times faster than the Opteron (2.2 GHz) based cluster.

Graph 2: Intel Itanium 2 outperforms Intel Xeon, which outperforms AMD Opteron

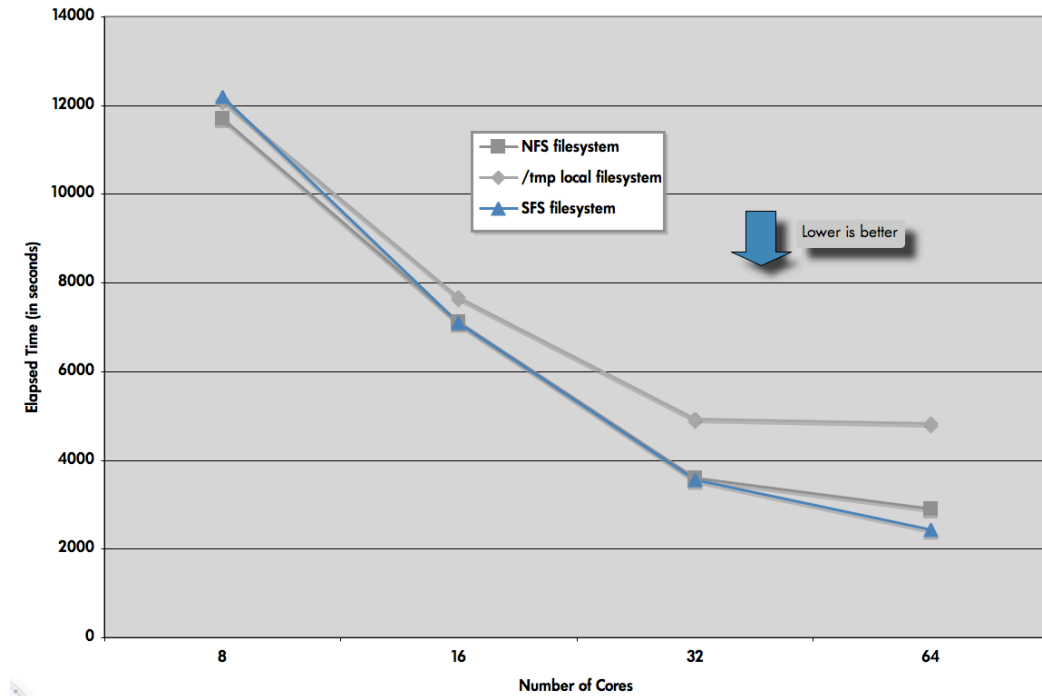


Configuration Details for Graph 2	
HP CP 6000	Architecture: Intel Itanium 2, 1.6 GHz Node: 2 processor 2 core rx1620 server Operating System: (Linux) XC v. 3.0 Interconnect: InfiniBand SDR
HP CP 3000	Architecture: Intel Xeon, 3.4 GHz Node: 2 processor 2 core HP ProLiant DL360 G4 server Operating System: (Linux) XC v. 3.0 Interconnect: Myrinet D
HP CP 4000	Architecture: AMD Opteron, 2.2 GHz Node: 2 processor 2 core HP ProLiant DL145 server Operating System: (Linux) XC v3.0 Interconnect: InfiniBand SDR

I/O and File System Access

I/O can often be a performance bottleneck when an application stores much of its data on disk. Network File System (NFS), the current de facto standard for accessing files, works well for small clusters of up to 16 cores, but I/O access time levels off as the cluster increases in size. Better performance is achieved when data is stored locally on a system's internal disks. HP StorageWorks SFS performs as well as NFS for smaller clusters, and it performs the best for larger clusters. This may be attributable to its ability to distribute files in parallel across clusters and storage components, resulting in higher aggregate bandwidth.

Graph 3: I/O Comparisons on a CP 6000 Cluster



Configuration Details for Graph 3

HP CP 6000

Architecture: Intel Itanium 2, 1.6 GHz
 Node: 2 processor 2 core rx1620 server
 Operating System: (Linux) XC v. 3.0
 Interconnect: InfiniBand SDR

HP StorageWorks SFS (over ELAN 4 interconnect): 3 OSS Nodes;
 12 Storage Enclosures (MSA 20)

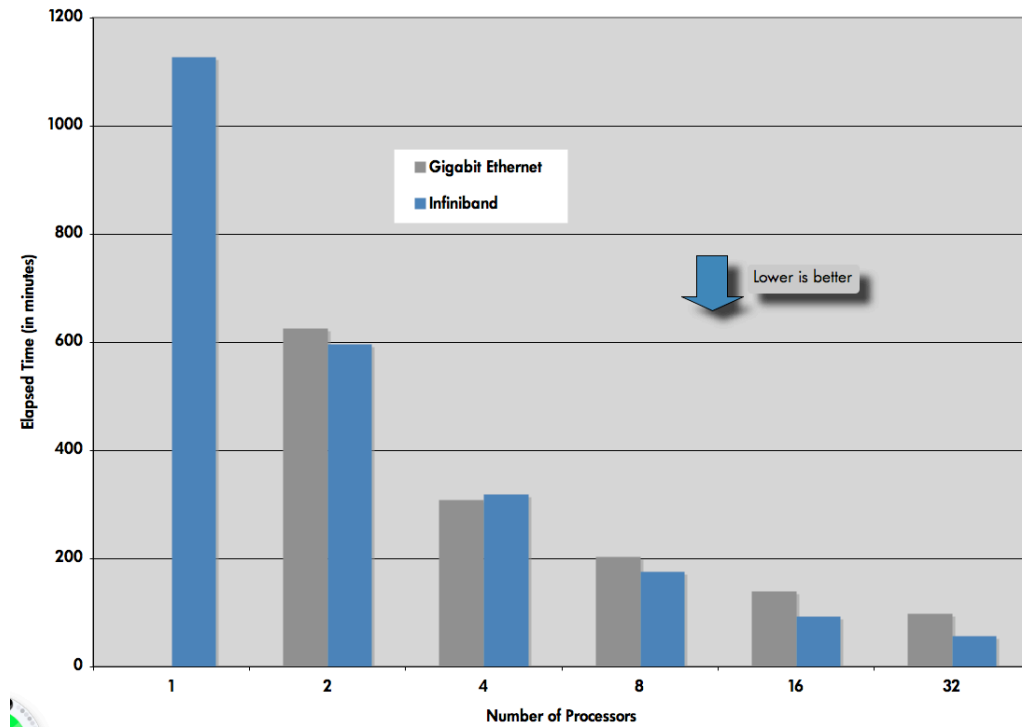
NFS (over ELAN 4 interconnect): 30 Drive RAID (0+1) Filesystem

/tmp: Single, Internal SCSI Drive (root)

Interconnects

With this test example, performance is comparable between InfiniBand SDR and Gigabit Ethernet for very small clusters (of up to 4 processors). With bigger clusters, job completion time is increasingly improved by Infiniband's higher speed (up to 750 MB/sec vs. GigE's 60-80 MB/sec) and lower MPI latency (<6 μ Sec vs. GigE's >40 μ Sec).

Graph 4: Interconnect Performance



Configuration Details for Graph 4

HP CP 6000

Architecture: Intel Itanium 2, 1.6 GHz

Node: 2 processor 2 core rx1620 server

Operating System: (Linux) XC v. 3.0

Disk: SFS using Gigabit Ethernet

Gigabit Ethernet and InfiniBand SDR Interconnect Tested

ADF has adopted HP-MPI for its cluster communications

Conclusion: Reducing Overall Computational Time

Clusters effectively reduce the overall computational time, demonstrating excellent scaling up to 128 processors. For DFT calculations, clusters of Itanium 2 servers, such as HP CP 6000, have a distinct performance advantage over similar clusters based on Xeon processors, which in turn outperform clusters based on Opteron processors. However, performance can be negatively affected even with the choice of the fastest computing architecture if memory and I/O are insufficient as ADF is a memory-intensive application and, for some job types, stores large amounts of data on disk. In this particular case study, each MPI process used a maximum of 700 MB of memory. Given today's typical configurations, memory was not a bottleneck to performance. I/O, on the other hand, can be improved from typical NFS installations. Adequate local disk is important for ADF jobs and for more than 16 cores, HP SFS delivers the best I/O performance.

HP Cluster Platform 6000 Performance

For DFT calculations, clusters of Itanium 2 servers, such as HP CP 6000, have a distinct performance advantage.

Finally, performance on clusters may be limited by interconnect speed. ADF spends little time communicating, so for clusters of 8 processors or fewer, the interconnect speed does not significantly affect overall performance. The interconnect speed becomes more important for larger clusters of 64 or 128 processors or more. Effective parallel speedups of large clusters require high performance interconnects. Because ADF supports HP-MPI, most interconnects are supported without a software recompilation. The choice of which interconnect to use depends more on whether inter-server communications will be a bottleneck in the mix of jobs to be run.

HP SFS Performance

Adequate local disk is important for ADF jobs and for more than 16 cores, HP SFS delivers the best I/O performance.

The parallel efficiency of a particular ADF job depends on many factors. If one uses a much more accurate numerical integration grid than the default, then the part of the code that is effectively parallelized moves towards 100%. Excellent scaling will be achieved for these simulations. The same is true for parts of the code which have not yet been optimized to the fullest, which increases the parallel percentage of the code and simultaneously removes several load balancing problems for serial parts of the code. Lengthy calculations that do not require frequent diagonalizations, such as analytic frequency calculations, are also among the best scaling ADF jobs. Finally, the symmetry of a molecule reduces both the size of the grid and the dimensions of the matrices that need to be diagonalized, which also affects parallel scaling.

HP continues to partner with commercial and academic software providers in Life & Materials Sciences. Collaborative projects to characterize key applications and optimize their performance on our computing solutions are a significant part of that partnership. For more information, check <http://www.hp.com/go/lifesciences>.

Glossary of HP Terms

CP	<p>Cluster Platform—HP’s cluster configuration families based on specific processor architectures:</p> <ul style="list-style-type: none"> • CP6000: HP Integrity server cluster family based on Intel Itanium 2 (2 processor/2 core [2p2c]) servers • CP3000: HP ProLiant DL server cluster family based on Intel Xeon (2p2c) servers • CP4000: HP ProLiant DL server cluster family based on AMD Opteron (4p4c) servers <p>HP provides consistent terminology to describe single core and dual core technology. The same terminology is used by both Intel and AMD.</p>
Number of processors	The number of physical chips (A processor may contain more than one core.)
Number of cores	The processing unit within a processor (There may be more than one core within a single processor/physical chip.)
HP-MPI	<p>HP’s implementation of the de-facto standard Message Passing Interface (MPI). HP-MPI supports multi-protocol execution of MPI applications on clusters of shared memory servers and provides transparent support of a broad range of interconnects and architectures. It supports HP-UX, Linux, and Tru64UNIX operating systems and interconnects such as TCP/IP, InfiniBand from Voltaire, Cisco, Silverstorm and Mellanox, Myrinet GM-2 and Myrinet MX, HyperFabric2 and Level 5 Etherfabric from Level 5 Networks. It also supports shared memory for intrahost transfers as well as Quadrics Ltd. QsNet on Linux Intel Itanium 2 based systems. HP-MPI is also thread-safe.</p>
HP SFS	<p>HP StorageWorks Scalable File Share. A high performance, high availability global I/O and shared file server that gives users of Linux clusters scalable storage (scaling to over 35 GB/sec aggregate bandwidth with a capacity that scales from 1 TB to 1024 TB) that is easy to use and easy to administer.</p>
UCP	<p>HP’s Unified Cluster Portfolio delivers flexibly configured clusters with a wide choice of platforms, systems, interconnects, and software options and the simplicity, reliability and value of a preconfigured, pre-tested, factory-built product.</p>
XC	<p>System Software supported by HP that is based on a standard Linux (Red Hat Enterprise Linux 4.0 Update 2) distribution combined with several open source packages.</p>

For more information

www.hp.com/go/lifesciences

© 2006 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Itanium is a trademark or registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

4AA0-8998ENW, November 2006

