

Performance Assessment of Multi-node Parallel Computing System from First-principles

^{1*}Mohammed B, ²Wante H.P, ³Hayatu A

^{1,2}Department of Science Laboratory Technology, Federal Polytechnic, Mubi, Adamawa State, Nigeria

³Department of Applied Physics, Federal Polytechnic, Mubi, Adamawa State, Nigeria

Abstract - A parallel computing environment of an interconnected set of computers called MPI Cluster is set up on a Linux Operating System to reduce runtime of Density Functional Theory DFT calculations by combining the computational power of multiple computers. In this paper, we evaluate the performance of Quantum ESPRESSO (QE) on the MPI Cluster system. To test the speed and scalability of our cluster system, distinct-point sample work loads are being distributed over multiple MPI processors. The result implies that scaling speedup over many processors is only possible if the number of k-points to parallelize is bigger than the number of processors. We also discovered that the speedup limit for parallelizing band calculations is somewhat independent of the number of bands employed and that it reduces linearly as the number of MPI processors increases.

Keywords: Density Functional Theory, MPI Cluster, MPI processors, parallel Computing, QE.

I. INTRODUCTION

Quantum ESPRESSO (QE) is a computer simulation package based on Density Functional Theory (DFT) for calculating electronic and structural properties of a material at ground state, which gives an excellent balance of accuracy and computational cost [1]. For a macro-molecular system with a large number of atoms, it takes several hours to execute even a simple calculation. The integration of a parallel library has made the package compatible to distribute work on many processors through the use of MPI. The computational cost is still challenging as a single computer have a limited number of processors.

The emergence of Density functional theory (DFT) [1] in 1968 has been a great achievement that contributed significantly to theoretical studies of electronic and structural properties of a material. It is an approach to describe the quantum behaviours of the atom in setting up a practical value. It is a well-established quantum mechanical method for electronic calculations for molecules. Principally, in electronic structure calculations, a molecular system is described by a set of functions that depend on the set of coordinates of all the particles in the system. This set of functions is known as a

wave function in quantum mechanics. The energy of the system is calculated from the wave function through quantum mechanical operations and from the solution of these wave functions. The complex nature of the wave function makes the evaluation of the full system very complex and highly computationally expensive even for the simplest molecules. The Hohenberg-Kohn [2] formalism of DFT approximates this ab-initio approach and proposes electronics density instead of electronics coordinates wave function. Electronic structure calculations were one of the main areas of application in high-performance computing (HPC) for the last decades. Over the past years, this method has rapidly grown as the cutting edge of the quantum mechanical theory that has been used by a large number of researchers in material science and other disciplines. While the methods used for these calculations have changed, the approach adopted by Car-Parrinello in 1985 [3] was one of the most common to be used. The Car-Parrinello method is a type of molecular dynamics, usually employing periodic boundary conditions, plane-wave basis sets, and density functional theory.

First-principles calculation within a framework of DFT is a renounced method for investigating material properties at the ground state. DFT in principle provides a true description of electronic behaviours and structural parameters at the atomic level of bulk material at ground states within some allowed approximation. The accuracy of calculation using the DFT approach is spanned within an error limit of 10-3 eV [4].

The complex nature of simulating solid material has led to the computational time remaining very high which pose a great challenge in computational condensed matter. Therefore, there is a growing demand for much more efficient implementations of DFT. Parallel computing involves the use of multiple processors (cores) of central processing units (CPU) of a computer to minimize these challenges. Researchers have been using parallel computing for several years when working on computationally difficult tasks such as simulation in high energy physics and bulk materials. The challenges continue as every computer have a limited number of processors that will available to contribute to parallel computing. A large complex problem requires hundreds or even thousands of processing hours to solve one problem and

which is not feasible to achieve on an ordinary computer. As such, the use of an interconnected set of computers called Cluster emerges which allows simultaneous use of many computer processors to solve one computational problem. Clustered computers are specifically designed to take a large set of data and divide them into parts and distribute them across the system entire processors thereby allowing the Cluster's node to process their unique task and finally collect them in their entirety.

In this work, we established a computational environment on a Linux operating system called MPI Cluster, where QE is being configured to use any number of processing resources available. The main aim of this paper is to investigate the distribution of parallelism over a locally connected cluster. We investigate the parallel scalability of QE over MPI-Cluster for performing a ground state total energy calculation; which is particularly important as often the total energy calculation needs to be performed many times during the geometry optimization of the atomic and electronic structure.

In condensed matter physics, the band theory of solids [5] allows a physical description of the occupied electron energy state in a crystalline solid. It enables us to explain the microscopic properties of the material, such as electrical, magnetic, thermal conductivity, optical absorption etc. One of the most important pursuits in solid-state fundamentals of the first-principles study of materials in condensed matter physics, the band theory of solids [5] allows a physical description of the occupied electron energy state in a crystalline solid. It enables us to explain the microscopic properties of the material, such as electrical, magnetic, thermal conductivity, optical absorption etc. One of the most important pursuits in solid-state physics is to discover new materials and study their properties to sustain continuous progress and improve the current electronic devices.

II. METHODOLOGY

A computer cluster is a group of linked computers that work together closely so that, in many aspects, they form a single computer. Our Linux cluster consists of a control node that serves as the main computer and four different computers that operate as slave nodes. Each computer in the cluster, including the master computer, is regarded as a cluster member or cluster node. The master computer communicates with other nodes (slaves) through a Local Area Network (LAN) connection. Quantum Espresso, the simulation software for this research is installed and configured in such a way that our computation can be executed on either the main computer (master node) or distributed across cluster available cores (mpi processors). The computing cluster has allowed QE to divide and distribute computational task or work-loads

across different processors through the application of message passing interface (MPIMPI is a standard message passing system which is used by QE to map out the total computational task and divide it into several chunks of data and distribute it to different processors accordingly. MPI does the intercommunication among different processes and combines the outcome from the various processes running on the different nodes to produce a unified outcome. Our Cluster system provides access to add new nodes (computers) to the system, and MPI gives the facility for the parallel running of QE to use those attached nodes.

2.1 Running of QE on cluster

QE runs directly from the master computer, which communicates with all slave nodes in the MPI Cluster. The Master computer is the only controlling node, which traces and communicates with only nodes (slaves) available within the MPI Cluster system, including itself. A computational task is performed on the master computer and it distributes the task across various processors through the use of MPI among the nodes (slaves). With the broad availability of multi-node mpi-processors, QE will be able to take advantage of the availability of several cores for most of its capabilities, be it ground-state SCF calculations, ground-state phonon calculations, molecular dynamics, etc.

2.2 System specification

The parallel benchmark run was conducted on a single machine with 12 threads of working processors (CPU), an Intel 3.6 GHz system with 31 GB of main memory per CPU. The other 4 nodes that form the MPI-cluster are interconnected with fast internet (LAN) network. We used mpich-3.3 as the MPI communication library. The source code of QE-6.4.1 was compiled with gfortran-4.8. The MPI-Cluster setup specification for each node can be referred to in Table 2.1.

Table 2.1: System specification for MPI-Cluster

Name (node)	Model(Intel-core)	Speed (GHz)	Physical Core	Memory	processors
Master	i7-4960x	3.6	6	32	12
Slave-1	i7-3770	3.4	4	8	8
Slave-2	i7-377x	3.5	4	8	8
Slave-3	i7-377k	3.5	4	32	8
Slave-4	i7-2880	2.27	4	12	8

2.2 The input file

All parallelized calculation has been executed with an essentially identical input file which defines a (rhombohedral) unit cell of Bi₂Te₃ containing two atoms of Bismuth and three atoms of Telluride as shown in Figure 2.1. We use plane wave cut off 45 Ry (resulting in 7440 plane waves) to expand the

Kohn-Sham orbitals for valence electrons. An ultra-soft (USP) pseudo potential of generalized gradient approximation (GGA) in the form of Perdew-Burke-Ernzerhof (PBE) [6] functional is used for exchange-correlation interaction between the electrons. A Monkhorst-Pack k-points mesh [7] was used to sample the first irreducible Brillouin zone of reciprocal space... All calculations are based on Density functional theory (DFT) within the plane-wave pseudo potential techniques and projected-augmented-wave (PAW) [8] as implemented in Quantum ESPRESSO simulation code.

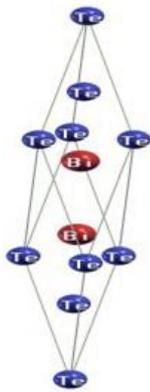


Figure 2.1: crystal structure of Bi2Te3

2.3 Computation method

QE offers parallelization (data distribution) over k-points, bands, and plane waves (G-vector in Fourier transformation from real to reciprocal space and vice versa). This division is controlled by the ‘mpirun-*np*’ command when executing the parallel calculation. For parallelization over k-point, the number of k-points along the axis of the Monkhorst-Pack mesh with different parameters of k-path which describe points along x, y, and z-axes of the grid point as shown in Table 3.1. For parallelization over bands, we use a fixed value of 721 k-points generated using the Xcrysden simulation package along with the high symmetry of the first Brillouin zone of the rhombohedral crystal structure [9] of Bi2Te3. The input files used for k-points and band parallelization are described in 2.2 above.

Table 2.2: System specification for MPI-Cluster

k-path along the axes of Monkhorst-Packmesh	Number k-points in irreducible BZ
4 × 4 × 4	13
5 × 5 × 5	19
6 × 6 × 6	32
7 × 7 × 7	44
8 × 8 × 8	65
9 × 9 × 9	85
10 × 10 × 10	116

2.4 The ground state self-consistency field calculation

First of all, we measured how long it takes for ground state energy execution and how it scales with the number of processors. The ground state SCF calculation is an iterative process that obtains the basic state of an atomic system. But before starting those iterations, some initialization is required, i.e. preparing the mpi environment, creating the mesh, reserving memory, etc., and also the generation of new charge density and wave function. The overall execution time is the sum of both parts (initialization + iteration + writing of the new wave function). And this is the essential step to starting any DFT and post DFT calculation. Other physical properties of the system can be determined from the wave function and charge density generated by SCF total ground state energy calculation.

III. RESULTS AND DISCUSSION

3.1 Performance of QE on cluster

The parallel performance of the MPI-Cluster system, which uses up to 40 threads (mpi processors), was analyzed by measuring the CPU time recorded for calculating the ground-state energy of Bi2Te3 at a different set of irreducible k-points. Table 3.2 contains the list of k-points generated by the Monkhorst-Pack k-points mesh [9] for this test. A similar calculation is done, but on a single computer having 12 processors, which serve as a benchmark for this test. We analyze parallel performance and scalability of the application of QE on a multi-node cluster system and describe the results with the knowledge gained about the application. The performance was evaluated by comparing the computational time recorded on the cluster system relative to those on a single computer for the same set of input parameters (input file). Table 3.1 summarizes the runtime for parallelism with different k-points executed on MPI-Cluster and single computers.

Table 3.1: Relative runtime for parallelism with different k- point executed on MPI-Cluster and Single computer

k-path along the axes of Monkhorst-Packmesh (nk)	Runtime(min)	
	Single computer	MPI Cluster
4×4×4	10.15	3.72
5×5×5	14.28	4.28
6×6×6	23.46	7.10
7×7×7	33.41	9.44
8× 8 × 8	46.53	14.39
9×9×9	64.00	19.00
10× 10×10	91.00	25.40

Referring to Figure 3.1, the k-point parallelization over MPI-Cluster gives us a small rise in computational time when

making calculations with a larger set of k-points as compared to a single computer, where we have a sharp rise in computing time (runtime) as the number of k-points is increased. For instance, using MPI-Cluster, we can calculate 10 10 10 k-point mesh in only 25 minutes, but at most, in 25 minutes you can only calculate 6 6 6 k-point mesh. Similarly, it will also take only 25 minutes to run 10 10 10 k-points in MPI-Cluster than on a single computer, where it would take not less than 90 minutes for the same number of k-points, and this would save more than 200% of computing time.

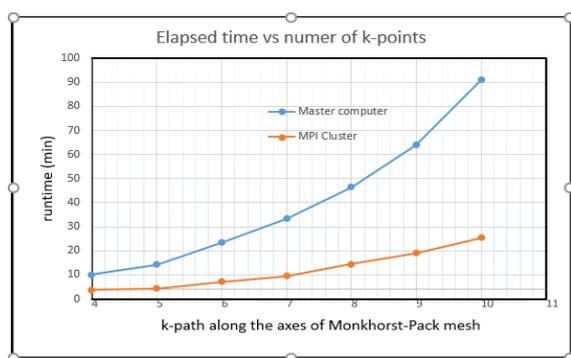


Figure 3.1: Relative timing for computing SCF of Bi2Te3 with different k-points on Single Computer and MPI-Cluster

3.2 Parallelization over k-points

Since the Eigen state at different k-points is almost independent of each other, interaction only occurs via density, it is natural to distribute the data over the k-points. This means we can distribute the data and workload efficiently using k-point parallelism. This type of parallelization can be very efficient because all three most time-consuming parts of the code (Hamiltonian matrix setup, diagonalization and generation of the new charge density) are independent for different k-points and there is no need to communicate during the calculation. That is why this type of parallelization is fine, even if the communication between the nodes/processors is slow. The effect of parallelization on k-points for this work is analyzed by timing total CPU runtime on computing ground state energy of Bi2Te3 for different number of processors with different k-points as presented in Table 3.2

Table 3.2: Parallelization of k-point over MPI-Cluster (time in minutes)

cores	32 k-points		65 k-points		116 k-point		189 k-point	
	Time	Speedup	Time	Speedup	Time	Speedup	Time	Speedup
1	26.5	---	54.07	---	89	---	157	---
10	7.8	3.4	13.55	3.99	24.25	3.67	39.34	3.99
20	5.7	4.65	10.22	5.29	19.05	4.88	29.23	5.37
30	4.69	5.1	9.51	5.68	17.45	5.1	28.39	5.53
40	5.2	5.1	8.8	6.14	16.82	5.29	27.54	5.7

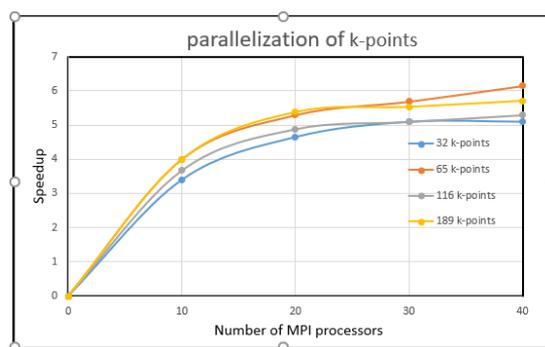


Figure 3.2: Speedup vs a number of processors of MPI-Cluster for parallelization over k-points in SCF ground state energy calculation

3.3 Parallelization over bands

Figure 3.2 Speedup vs a number of processors of MPI-Cluster for parallelization over k-points in SCF ground state energy calculation.

Parallelization over Kohn-Sham orbitals (i.e., bands) becomes favourable when the number of bands is so large that it begins to dominate in terms of computational time. Linear algebra for orthonormalization and diagonalization of the wave-functions is the most noticeable contributor to this regime, and therefore, band parallelization can be used to distribute the computational load over several processors of a computer. The Kohn-Sham states are split across the processors of the band group. Some calculations can be independently performed for different band indexes.

Table 3.3: Parallelization of bands over MPI-Cluster (time in minutes)

cores	30 bands		40 bands		50 bands		60 bands	
	Time	Speedup	Time	Speedup	Time	Speedup	Time	Speedup
1	120	---	190	---	215	---	271	---
10	55.29	2.17	85.20	2.23	95.13	2.26	129.04	2.10
20	31.49	3.81	50.39	3.77	55.26	3.89	75.07	3.61
30	27.08	4.43	43.37	4.38	50.59	4.25	67.07	4.04
40	24.14	4.97	40.33	4.71	43.93	4.93	59.96	4.52

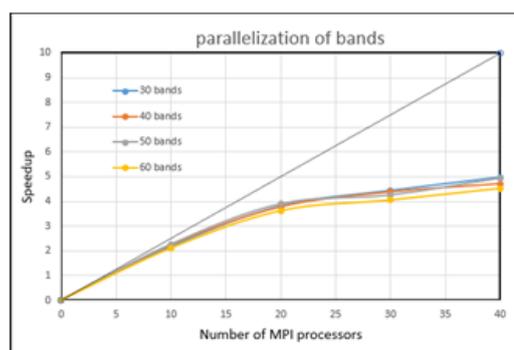


Figure 3.3: Speedup vs a number of processors of MPI-Cluster for parallelization over bands

3.4 Speedup analysis

The speedup is defined as the time taken for sequential calculation divided by the time taken for the parallel calculation. Figures 3.2 and 3.3 show the speed up versus the number of processors on the MPI-Cluster for parallelization over k-points and bands respectively. We measure the speedup of the MPI-cluster by parallelizing 5 sets of k-points, namely 32, 65, 116, and 189, each over a different set of mpi-processors. From Figure 3.2, we observed almost identical speedup for parallelization over a different set of k-points up to 30 processors, except at a point when the number of cores exceeded the number of k-points to be parallelized. A decline in speed-up over 30 cores is observed for 32 k-points (Figure 3.2, blue line). This happens because the 32 k-points need to be distributed over a max of 32 processors. Saturation of speedup always occurs when the number of processors exceeds the number of k-points. Ideally, if QE is given more processors than it has k-points for a calculation, then it will further distribute the data and workload over the plane-waves (G-vectors'), which affects the overall performance. This parallelism works well for small numbers of processors but requires more communication between the processes as the number of processors increases, which may eventually result in poor scalability. For band analysis, we ran a series of calculations with a different set of bands each parallelized over 10, 20, 30, and 40 processors as shown in Table 3.3 and is limited to a fixed number of k-points. The result from Figure 3.3 shows that parallelization of bands shares similar scaling for the different numbers of bands used and there is a fast increment in the speedup until 20 processors. The speedup seems to be very slow, between 20 to 40 processors, with an average increment of 1. Therefore, we have found the limit of speedup for parallelization of band calculations is partially independent of the number of bands used and it decreases as the number of cores increases.

IV. CONCLUSION

Computer simulation is one of the most powerful techniques scientists use to understand any kind of physical phenomenon that might be difficult or rather impossible to explore experimentally. Supercomputer facilities offer the possibility of using hundreds of thousands of computing processors to accelerate the execution of those simulations, giving the opportunity to investigate the behavior of a very complex system. This dissertation is concerned with the use of a multi-node parallel computer to simulate materials based on the Density Functional Theory approach. The study is centered on an open-source simulation package, Quantum Espresso, although they obtained results are not limited to only this package. The package can run on a multi-node computing system if properly configured, via the multi-thread MPI

library. The QE was developed for DFT calculation and is used to study first-principles the properties of the ground states of a crystalline solid or molecular system.

We have tested parallelism over k-point and bands in QE within the MPI-Cluster system. The result suggests that the improvement in scaling over many processors is limited only if the number of k-points to be parallelized is greater than the number of processors. Parallelization of k-points is more advantageous for a small system (small unit cell) demanding a lot of k-points. The k-point parallelization will give us increased speed when making calculations with large k-point sets. We also found the limit of speedup for parallelization of band calculations is partially independent of the number of bands used and it decreases as the number of processors increases.

ACKNOWLEDGEMENT

We acknowledged University Technology Malaysia, computer laboratory for the wonderful collaboration by providing a conducive and enabling environment that made it possible to carry out this research work.

REFERENCES

- [1] R.G.Parr, "Density Functional Theory in Chemistry," in Density Functional Methods In Physics, Boston, MA: Springer US, 1985, pp. 141–158.
- [2] I.E. Gaa, P. Hohenberg, X. Superzeure, F. And, and W. Konnt, "PHYSICAL REVIEW." Accessed: Jun. 21, 2019. [Online]. Available: <http://users.wfu.edu/natalie/s15phy752/lecturenote/HohenbergPhysRev.136.B864.pdf>.
- [3] P.Giannozzi et al., "QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials," J. Phys. Condens. Matter, vol. 21, no. 39, pp. 9–11, 2009, doi: 10.1088/0953-8984/21/39/395502.
- [4] A.Lawal and A. Shaari, "Density functional theory study of electronic properties of Bi₂Se₃ and Bi₂Te₃," Malaysian J. Fundam. Appl. Sci., vol. 12, no. 3, Jan. 2017, doi: 10.11113/mjfas.v12n3.424.
- [5] B. Wilkinson and M. Allen, "PARALLEL PROGRAMMING TECHNIQUES AND APPLICATIONS USING NETWORKED WORKSTATIONS AND PARALLEL COMPUTERS 2nd Edition," 2015.
- [6] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," Phys. Rev. Lett., vol. 77, no. 18, pp. 3865–3868, Oct. 1996, doi: 10.1103/PhysRevLett.77.3865.
- [7] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," Phys. Rev. B, vol. 13, no.

12, pp. 5188–5192, Jun. 1976, doi: 10.1103/PhysRevB.13.5188.

- [8] P. E. Blöchl, “Projector augmented-wave method,” *Phys. Rev. B*, vol. 50, no. 24, pp. 17953–17979, Dec. 1994, doi: 10.1103/PhysRevB.50.17953.
- [9] X. Gao, M. Zhou, Y. Cheng, and G. Ji, “First-principles study of structural, elastic, electronic and thermodynamic properties of topological insulator Bi₂Se₃ under pressure,” *Philos. Mag.*, vol. 96, no. 2, pp. 208–222, Jan. 2019, doi: 10.1080/14786435.2015.1128126.



Name: **Wante Haruna Peter**
Country: Nigeria
State: Adamawa
City: Mubi
Occupation: Lecturer
Specialization: Energy Physics



Surname: **Hayatu Abubakar**
Country: Nigeria
State: Adamawa
City: Mubi
Department: Applied physics
Institution: Federal Polytechnic, Mubi
Specialization: Applied Physics

AUTHORS BIOGRAPHY



Name: **Mohammed Buba**
Country: Nigeria
State: Adamawa
City: Mubi
Occupation: Lecturer
Specialization: Solid State Physics

Citation of this Article:

Mohammed B, Wante H.P, Hayatu A, “Performance Assessment of Multi-node Parallel Computing System from First-principles” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 5, Issue 8, pp 82-87, August 2021. Article DOI <https://doi.org/10.47001/IRJIET/2021.508014>
