Role of Parallel Computing in Numerical Weather Forecasting Models

Subhendu Maity[†], SubbaReddy Bonthu, Kaushik Sasmal, Hari Warrior
[†] National Institute of Technology Meghalaya
Shillong – 793003
Indian Institute of Technology Kharagpur,
Kharagpur – 721302

ABSTRACT

Parallel computing plays a crucial role in state-of-the-art numerical weather and ocean forecasting models like WRF, POM, ROMS and RCAOM. The present study is an attempt to explore and examine the computational time required for the highly complex numerical simulations of weather and ocean models with multi core processors and variable RAM/processor speeds. The simulations, carried out using machines of different computational capability/configuration viz. quad core and Xeon machines, have been investigated with different synthetic experiments to evaluate the role of parallel computing in the operational forecasting system. The saturation rates with different number of processors are also calculated before carrying out forecasting studies. Serial and parallel computations have been carried out with WRF (Weather Forecasting Model) model for simulating the track of a natural hazard viz. the Thane cyclone. The simulations reveal that in the initial stage the computational time decreases exponentially with number of processors and later it reaches saturation stage, even though the number of processors is increased. Additionally, parallel computing simulations showed that the model simulations depend upon the model time step, grid resolution, number of cells in the domain, system architecture, and finally number of vertical levels and their resolutions.

Keywords

parallel computing, grid resolution, quad core machine, atmospheric model, ocean model, Xeon Machine.

1. INTRODUCTION

Numerical weather prediction is an application of numerical algorithm and it is used to predict the future state of atmosphere by applying the initial state of weather [1]. In 1960, vector and vector shared memory computers have been used for the numerical simulations by offering the performance first in the tens, then hundreds, then thousands of megaflops. The increasing computer power leads to finer scales, longer simulations, and increasing sophistication in the modeling of atmospheric and oceanic processes. The parallel computers are more sophisticated to provide potential performance in the hundreds of gigaflops memory capacity to allow problem sizes of previously intractable dimensions. The parallel computers showed their role to give real-time forecasts of individual features of a severe storm and multicentury simulations of climate like ENSO and global warming. In 1965, John Drake and Ian Foster edited the special issue of Parallel Computing on climate and weather modeling. It explored the issues, which were related to adaptation of global models of the earth's atmosphere and oceans to parallel computers. The special issue was focused on the special characteristics, problems, and implementation strategies of parallel computing for a closely related branch of geophysical applications viz. regional scale models.

The regional weather and ocean modeling are important branches of atmosphere and ocean science, with application to forecasting, climate change study and basic atmosphere and ocean research. These models are playing a crucial role to get detailed simulations of particular area of interest which can not be achieved through the Global models. The regional models are called 'limited-area' models. These models are closely related historically, methodologically, and operationally to global models. They solve the basic equations that describe atmospheric motion, though the formulations may differ, and also they use similar modules to simulate atmospheric processes such as clouds and radiation. These models are used as counterpart to global models, having a number of important structural and algorithmic differences that bear on their suitability of parallel computers. Initially parallel computing was focused on computing of climate and weather changes by using global scale models. The computational cost of an atmospheric model is a function of the number of cells in the domain and the time step. Explicit three dimensional hydrodynamic codes typically cost O(n4) calculations, where n is a grid dimension. The fourth-order term reflects the additional refinement of the time dimension (smaller time -step) necessary to maintain numeric stability as the spatial dimensions of each cell are decreased. Because the atmosphere is shallow, and because the number of vertical layers is independent from and more or less constant with respect to the horizontal resolution, the cost behaves more are like a function of n3. Even so, a doubling of resolution (a half of the width of a cell) results in an eight-fold increase in computational cost. Therefore, resolution is a precious commodity, and there is a strong incentive not to waste it. Regional models permit computational effort to be concentrated over a limited area for the purpose of computing higher-resolution simulations than would be possible globally. This capability is useful, for example, for simulating the effects of complex terrain on an evolving weather system. Nesting, the ability to refine the mesh over sub-domains within a regional simulation, extends the ability of some regional models to efficiently concentrate on resolution. In terms of numerical methods, regional models are simpler than global models, employing finite-difference methods without special treatment required around a pole. The use of finite difference methods also has favorable implications for parallelization, since such methods are nearest-neighbor and therefore more straightforward to implement efficiently than the spectral method employed in some global models.

2. PARALLELIZATION

The regional weather forecasting models include the components of dynamics and physics. The dynamics of model takes care about model integration and the physics component of a regional model includes physical processes such as long- and short-wave radiation, condensation and precipitation, and surface processes. The sub-grid scale processes are modeled by parameterization rather than differential equations.

The atmospheric model and its applications implementation on a parallel computer involves partitioning the work among processors so that they can work together to perform the computation faster than the work when it is performed sequentially. Atmospheric models carry the same set of computations in each vertical column of their three-dimensional, regular, and almost universally Cartesian domains. The SPMD (single program, multiple data stream) method is used in one or more dimensions of the domain to divide it into sub-domains and make rectangular tiles, and finally allocate to separate processors. In a distributed-memory machine, the state and intermediate data representing the sub-domain are transferred to the local memory of a processor, where the data are readily accessible by that processor but accessible to other processors only through inter-memory transfer: explicit message passing, shared-memory “gets” and “puts,” or implicit data movement through cache-coherency hardware and software. The cost of data movement is of efficient concern because, unmitigated, it is pure overhead.

The model physics sends information to the processor decomposition in atmospheric models, because the data dependency is more important in the vertical directions rather than in the horizontal. The code is made parallel by decomposing over the two horizontal dimensions, viz. latitude and longitude; communication is removed within physics, and this makes the code almost perfectly parallel. It is desirable from the software perspective to use existing sequential physics packages without modifications, since these are the parts most liable to change over the life of the model code [2].

The explicit Eulerian finite-difference schemes communicate by exchanging data between edges of neighboring sub-domains and the edge exchange width is fixed. The other schemes like semi-Lagrangian schemes also communicates data between neighboring sub-domains, but the edge exchange width is a dynamic function of model state. The fixed overlapping region will be larger than Eulerian schemes. The implicit methods, in which the value computed at a point depends not only on its neighbors but on the values of all the other points in the domain. Direct methods employ parallel data transposition; parallel iterative solvers use global summation along with pre-conditioners. The pre-conditioners may also entail communication, or they may be strictly local algorithms.

The parallel models suffered with communication cost, load imbalance. The load imbalance comes because of the fact thst some processors have more work than others (or if some processors are more powerful than others), the faster processors will reach a common synchronization point (communication or I/O) and then be idle until the slower processors catch up. It includes uneven distribution of points to processors in a dimension, different amounts of work at domain boundaries, different amounts of work having to do with the localized state of the atmosphere in part of the domain (physics imbalance), unequal processor speeds, and unequal loading of processors from other jobs if the model does not have exclusive access to the nodes. Imbalances can be addressed by reallocation of work to processors, either statically at the beginning of a run or dynamically over the courses of a run; frequently, however, such imbalances are simply ignored. The last option is not unreasonable if the cost or complexity of addressing an imbalance outweighs or marginalizes the benefit; nevertheless, several of the papers in this issue quantify and attempt to remedy load imbalance, with some success.

Now many models have some form of mesh refinement for concentrating resolution over a particular feature of interest in a simulation. In one-way nesting where the lateral boundary conditions are forced to the nest from the parent domain but without feedback, is computationally the simplest mesh refinement strategy, because forcing is infrequent (on the order hours between each new set of lateral boundary conditions) and the parent and nest can run asynchronously in batch mode. Two-way nesting is useful by getting the feedback and has the advantage of keeping the parent and nest solutions from diverging into separate trajectories, but it is computationally more demanding. In the two-way nesting, the forcing and feedback occur at every time step, and the parent and nest must run synchronously. Parallelization of such schemes is also more complex, since it requires a choice of costly communication between parent and nest running in parallel, suffering load imbalance on the parent in order to avoid communication forcing and feedback data between processors.

3. DESCRIPTION OF MODEL – PARALLEL COMPUTING

In the present study, the advanced Weather Research Forecast (WRF) mesoscale model version 3.2 is used to measure the intensity of maximum sustainable winds and track prediction of the Thane cyclone. WRF model is widely accepted community model in the field of atmospheric science and is a next generation mesoscale model developed by National Centre for Atmospheric Research (NCAR) Mesoscale and Microscale Meteorology (MMM) division in collaboration with number of research and operational organizations of USA. Mchalakes et al. and Skamarock et al. [3] reported that WRF model is non-hydrostatic mesoscale model which is used for the simulation and prediction of fine-scale atmospheric phenomenon emphasized with horizontal grid resolutions ranging from hundreds of kilometers to meters. It is based on the Eulerian solver for fully compressible non-hydrostatic equations, cast in flux conservation form (Ooyama [4]) using (hydrostatic pressure) Eulerian mass dynamical core with terrain following hybrid sigma-pressure vertical coordinates (Laprise [5]). Dudhia [6] and Skamarock et al. [3] explained the detail description of WRF model including equations, physics and dynamics of the model. The simulations of meteorological events at different scales have been well tested using logical combination of different physical parameterization schemes and data assimilation techniques. In the present study, the model grid system is supported with staggered Arakawa-C grid with an integration time step of 150 sec. The model simulations have been integrated with 2nd and 3rd order Runga-Kutta time integration scheme and the horizontal spatial distribution has been estimated with the 6th order centre differencing scheme. The second to sixth order advection schemes has been used in horizontal and vertical directions. It has got many parameterizations of physical processes, user options, which are growing interest to bring the model into broad modeling community. It possesses a number of physics schemes, initialization routines, and data assimilation packages. The available convective parameterization schemes in the model are Kain-Fritsch new Eta scheme (Kain and Fritsch [7]), Betts-Miller-Janjic scheme (Betts and Miller [8], Janjic [9]) and Grell-

Devenyi ensemble scheme (Grell and Devenyi [10]). In the present case study, The Yonsei university (YSU) boundary layer scheme (Hong et al. [11]) is used in WRF modeling system. The most regular used microphysics schemes are Lin et al scheme, Ferrier (new Eta) microphysics, and WSM-6 class Graupel scheme to perform the experiments in the Bay of Bengal. The complete details of model configuration of present study for the simulation of Thane cyclone is illustrated in Table-1. In the present study various schemes of convective parameterization and cloud microphysics schemes are employed. Dudhia short wave radiation scheme is used in all the experiments carried out in this study.

The WRF model equations are formulated using a terrain-following hydrostatic-pressure vertical co-ordinate and is denoted by η

$$\eta = (\mathrm{p}_h - \mathrm{p}_{ht}) / \mu$$

where

$$\mu = \mathrm{p}_{hs} - \mathrm{p}_{ht} \qquad (1)$$

p_h is the hydrostatic component of the pressure, and p_{hs} and p_{ht} are representing the pressure at surface and top boundaries. μ represents the mass per unit area within the column in model domain.

The flux form variables are

$$V = \mu v = (U, V, W)$$
$$\Omega = \mu \dot{\eta}$$
$$\Theta = \mu \theta \qquad (2)$$

$v = (u, v, w)$ are the covariant velocities in the two horizontal and vertical directions, respectively. $w = \dot{\eta}$ is the contravariant 'vertical' velocity, θ is the potential temperature. The WRF model is also having non-conserved variables like $\phi = gz$ (the geopotential), p (pressure), and $\alpha = \frac{1}{\rho}$ (the inverse of density).

The prognostic equations of WRF model in the Eulerian flux form are given as

$$\partial_t U + (\nabla . V_u) - \partial_x (p\phi_n) + \partial_n (p\phi_x) = F_U \qquad (3)$$
$$\partial_t V + (\nabla . Vv) - \partial_y (p\phi_n) + \partial_n (p\phi y) = F_V \qquad (4)$$
$$\partial_t W + (\nabla . V_w) - \partial_y (p\phi_n) - g(\partial_n p - \mu) = F_W \qquad (5)$$
$$\partial_t \Theta + (\nabla . V\theta) = F_\Theta \qquad (6)$$
$$\partial_t \mu + (\nabla . V) = 0 \qquad (7)$$
$$\partial_t \phi + \mu^{-1}[(V . \nabla_\phi) - gW] = 0 \qquad (8)$$

The diagnostic relation for the inverse density and equation of state are given as

$$\partial_\eta \phi = -\alpha\mu \qquad (9)$$
$$\mathrm{p} = \mathrm{p}_0 (R_d \theta / p_0 \alpha)^\gamma \qquad (10)$$

where $\gamma = {c_p}/{c_v} = 1.4$ is the ratio of heat capacities for dry air, R_d is the gas constant for dry air and p_0 is the reference pressure. The terms F_U, F_V, F_W and F_Θ represents the forcing terms arising from model physics, turbulent mixing, spherical projections and the earth's rotation. WRF model has been simulated for entire period of *Thane* cyclone to predict the best model track and accurate land fall time with minimum displacement error from Thane cyclone track. The sensitivity of the track prediction and land fall time of the cyclone has been calculated by using different model experiments with changing of initial and lateral boundary conditions.

4. MODEL CONFIGURATION – ITS EXPERIMENTS

The WRF model has been configured for Bay of Bengal to simulate the Thane cyclone with two nested domains having horizontal resolution of 27 Km and 9 Km with 27 pressure levels and their corresponding grid points are 162× 162 and 259× 259 respectively (Table-I). The centre of model domain is located at 80°E and 15°N. The model topography is obtained from the USGS at 2' and 30' resolutions respectively. The model simulations have been carried out by providing the initial and boundary conditions from the National Centre for Environmental Prediction (NCEP) FiNaL Analysis (FNL; 1°×1°) and Global Forecasting System (GFS) data. The time varying lateral boundary conditions are derived at every 6 h interval from the NCEP (GFS) forecast fields. The estimated/observed intensity and the position of the tropical cyclone are obtained from Joint Typhoon Warning Centre (JTWC) and Indian Meteorological Department (IMD) for comparison with model simulations. In the present study, a series of sensitive experiments are carried out for the Thane cyclone simulations of track, intensity and time of land fall to critically evaluate the WRF model performance with number of initial conditions. The performance of model has been carried out with different initial conditions using the logical combination of different parameterization schemes of microphysics and cumulus convection. First the model has been initialized on 26th December 2011 0000UTC and integrated up to 96 h i.e. 30th December 2011 0000UTC and performed simulations with all combination of schemes of cumulus convection and microphysics schemes and after each scheme of cumulus convection and microphysics were used with different initialized conditions.

5. RESULTS AND DISCUSSIONS

The regional atmospheric and ocean models are composed of numerical equations containing dynamics and physics. The fluid motions are calculated by using the integration of differential equations of dynamics and with concerned physics in the form of forcing terms and depend upon the model state. The algorithms are chosen to calculate the forcing term with the choice of physics options and implemented with concern for computational cost, stability, and accuracy. The choice of algorithms has been used as implications for the parallelization issues like data decomposition, communication, and load balance.

The dynamics of the model depends on the advection and gravity pressure balance and moreover based on the system of nonlinear time-dependent partial differential equations known as the primitive equations. The system of equations helps to calculate and carry the information of the computed fields, such as wind, temperature, pressure, and humidity. Dynamics of model has taken care of the approximations like a hydrostatic approximation, which assumes that the force of gravity is balanced by the vertical components of the pressure gradient, or a non-hydrostatic formulation. The stability of the model is maintained by a high order diffusion operator.

The Eulerian time integration scheme is an explicit scheme, which is used to solve the system of equations. In this scheme, the computation of a new time step is based on values from previous time steps only. Otherwise, the schemes are known as implicit. For an explicit scheme, the time step is limited by the fastest moving waves, which in a hydrostatic model, are gravity waves. To increase the time step, the terms responsible for the fastest moving waves may be treated by an implicit scheme. Such schemes are referred to as semi-implicit. Alternative, the terms responsible for the fastest moving waves may be integrated with a smaller time step than the remaining terms.

The model has been set up covering the Bay of Bengal for the simulation of the Thane Cyclone (Fig. 2). The model has been simulated for 96 h by using the initial condition of 26/00UTC and concerned physics options by the logical combination of cumulus parameterizations. The best track has been estimated by simulating WRF model using GD Ensemble cumulus Scheme (Fig.3) and it has been compared with the available observational track data available from IMD and JTWC. The predicted track shows clearly good comparison with observational track data sets, but it shows the leading time of land fall position is crossed 6h early than the observational land fall time. The disadvantage of this cumulus scheme for this cyclone is observed that the intensification of the cyclone is not simulated well as the other cumulus schemes do.

The model simulations have tested in both the serial as well as parallel computations. The WRF model has been simulated for the Thane Cyclone with Quad core machine (4 processors and 8 GB RAM) and Xenon machine (8 processors and 16 GB RAM) for 10h. The model simulated time has been evaluated in accordance with number of processors. The model simulation time has shown maximum time taken (520 min) in serial mode with Quad core machine and after that simulated time gradually decreases with increasing number of processors (Fig.4). The model simulated time with same configuration and same resolution but using different parallel computing system having 8 processors and 16 GB RAM show that the time taken is very less (160 min) in serial mode compared to that of Quad core machine simulations in the serial mode and later on its simulation time decreases gradually. The model simulations have revealed that the simulation time decreases up to certain number of processors and after that the model has reached to its saturation state which means that the simulated time may not decrease even though the numbers of processors is increased (Fig.5).

The comparative study made between these two parallel computers revealed that the computational speed depends upon the computational algorithms and system memory locations and finally its RAM/processor speed. The Quad core machine reveals that it takes more computational time in terms of numerical simulations for weather forecasting as it has less memory and RAM speed comparative to the Xenon machine, which has large memory and 16 GB RAM speed. In the sequential simulations both the machines have shown high simulation time but the results are clearly dependent on system algorithms and their memory capacity and RAM speed (Fig.6). In the parallel computations, initially the model simulation time has decreased exponentially up to certain number of processors and after that there is no significant time variations found.

6. CONCLUSIONS

The present study is aimed to examine and explain the role of parallel computing in the weather forecasting models. The study explains the variations in the model simulation time in the sequential mode as well as the parallel mode. The model simulations have taken huge time in sequential mode and less time in the parallel mode. Even in the sequential mode, different system algorithms and memory locations and RAM speed have shown their effect in the variation of the simulation time. The results confirm that the parallel computations are very helpful for making rapid simulations with more number of processors for quick and effective weather forecasting. In the computations it was revealed that the simulation time increases up to certain number of processors after which the variation of simulation time is comparably less with increase in the number of processors.

TABEL I: WRF DETAILS

Dynamics	Non hydrostatic
Data	NCEP FNL; GFS
Interval	6 hrs, 3 hrs
Grid size	Domain 1: (162× 162) × 27; Domain 2: (259× 259) × 27
Resolution	Domain 1: 27 km × 27 km; Domain 2: 9 km × 9 km
Covered area	2º-25 º N and 72º-97º E
Map Projection	Mercator
Horizontal grid system	Arakawa-C grid
Integration time step	150 sec
Vertical coordinates	Terrain-following hydrostatic pressure vertical co-ordinate with 51 vertical levels
Time integration scheme	3rd order Runga-Kutta Scheme
Spatial differencing scheme	6th order center differencing
PBL Scheme	YSU
Surface layer Parameterization	Noah land Surface Scheme
Microphysics	1. Ferrier (new Eta) 2. WSM 6-class graupel scheme
Short wave radiation	Dudhia scheme
Long wave radiation	RRT
Cu-physics	1. Kain-Fritsch (new Eta) scheme 2. Grell-Devenyi Ensemble

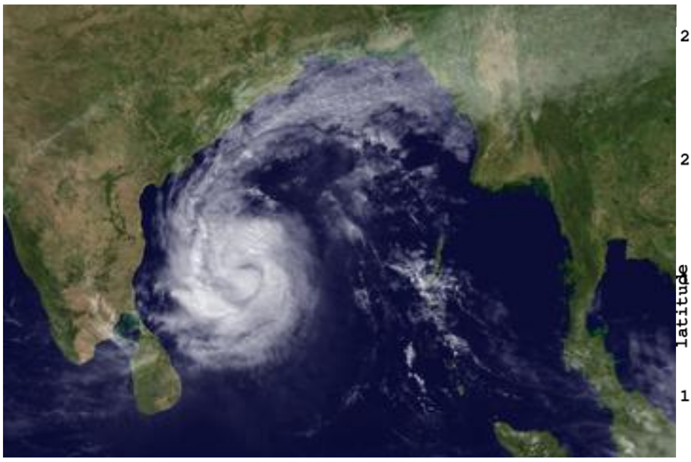

Figure 1 Typical satellite imagery of very severe cyclonic storm, *THANE* eye over the Bay of Bengal

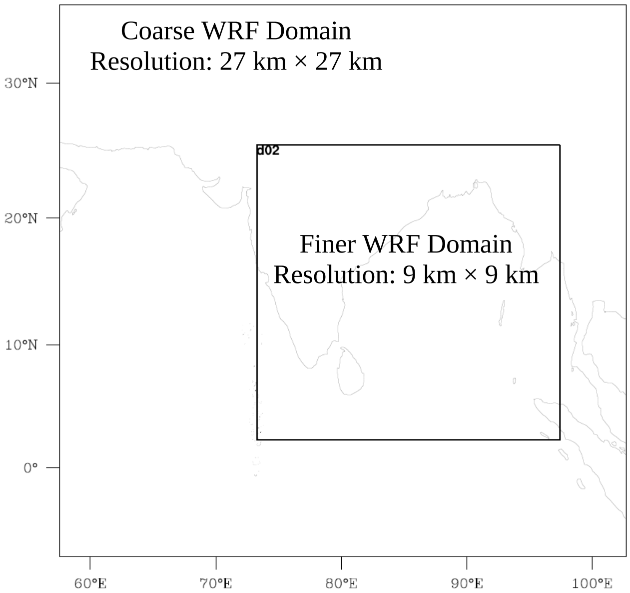

Figure 2 The domain configuration of WRF model for the *Thane Cyclone* simulations

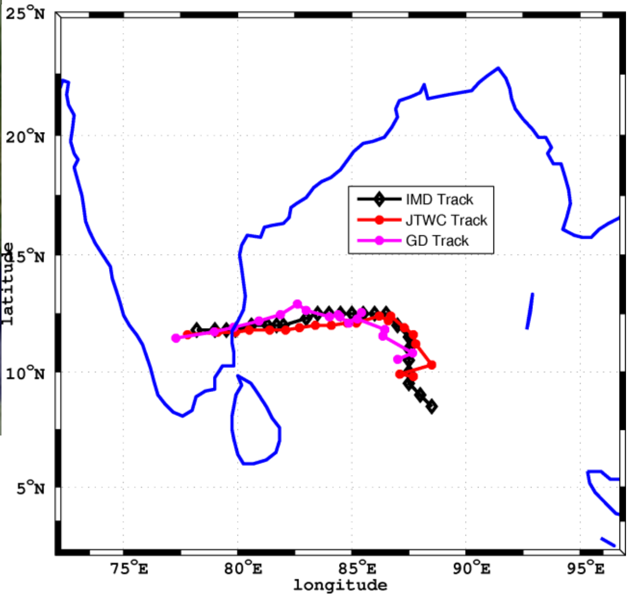

Figure 3 comparison of WRF model simulated track (by using GD Ensemble cumulus scheme) with IMD and JTWC Track

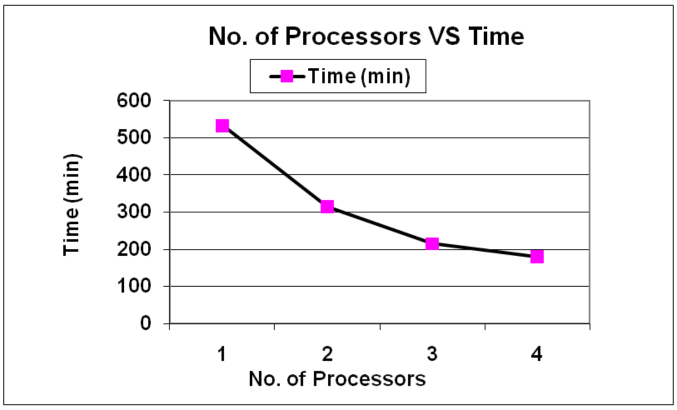

Figure 4 WRF model simulation time corresponding to the number of processors by using the Quad core machine

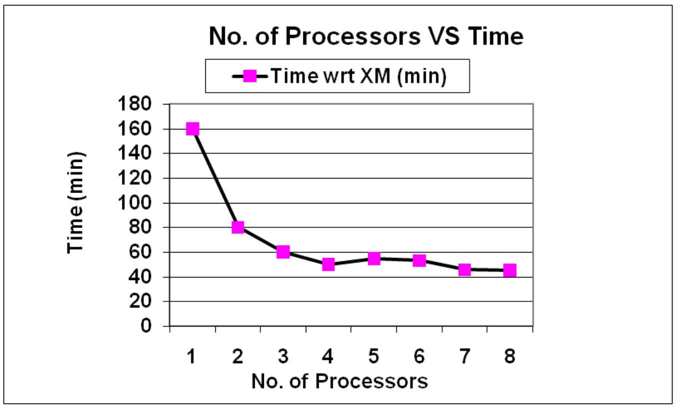

Figure 5 WRF model simulation time corresponding to the number of processors by using the Xenon machine (8 processors with 16GB RAM)

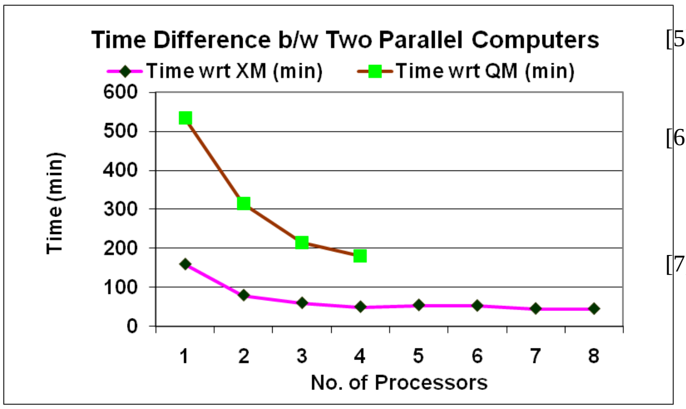

Figure 6 The comparison of WRF model simulation time with Quad core (4 processors with 8GB RAM) and Xenon machine (8 processors with 16 GB RAM)

7. REFERENCES

[1] R. Hardy, Weather, NTC Publishing Group, 1996.

[2] I. Foster, Designing and bulding parallel programs, Addison-Wesley, 1995.

[3] W. C. Skamarock, J. B. Kelmp, J. Dudhia, D. O. Gill, D.M. Barker, W. Wang and J. G. Powers, A description of the advanced research WRF version 2, NCAR Technical Note, 2005.

[4] K. V. Ooyama, A thermodynamic foundation for modeling the moist atmosphere, J. Atmos. Sci., Vol.47, 1990, pp. 2580-2593.

[5] R. Laprise, The Euler equations of motion with hydrostatic pressure as independent variable, Mon. Wea. Rev., Vol. 120, 1992, pp. 197-207.

[6] J. Dudhia, The weather research and forecasting model (Version 2.0) 2nd international workshop on next generation NWP model, Yonsei University Seoul, Korea, pp. 19-23, 2004.

[7] J.S. Kain, and J. M. Fritsch, Convective parameterization for mesoscale models: The Kain-Fritcsh scheme, The representation of cumulus convection in numerical models, K.A. Emanuel and D. J. Raymond, Eds., Amer. Meteor. Soc., pp. 246, 1993.

[8] A. K. Betts, and M.J. Miller, A new convective adjustment scheme. Part II: single column tests using GATE wave, BOMEX, and arctic air-mass data sets, Quart. J. Roy. Meteor. Soc., Vol.112, 1986, pp.693-709.

[9] Z.I. Janjic, The step-mountain eta coordinate model: further developments of the convection, viscous sublayer and turbulence closure schemes, Mon. Wea. Rev., Vol. 122, pp. 927-945.

[10] G. A. Grell and D. Devenyi, A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophy. Res. Lett., Vol. 29(**14**), 2002, Article 1693.

[11] S.-Y. Hong and J.-O.J. Lim, The WRF Single-Moment 6-Class Microphysics scheme (WSM6), J. Korean Meteor. Soc., Vol. 42, 2006, pp.129-151.