Benchmark of WRF-CHEM Model on HPC Cluster*

Shukla Foram P
Department of Masters in Computer Engineering,LJIET
Gujarat Technological University
Ahmedabad, India
shuklaforamp@gmail.com

Dr. Varun Sheel
Associate Professor,
Space&Atmospheric Science Division
Physical Research Laboratory
Ahmedabad, India
varun@prl.res.in

***Abstract*—The WRF-CHEM is a Weather research forecasting model for chemistry. WRF-CHEM model is widely used for atmospheric research institutes for Prediction and simulation of weather, or regional and local climate. This model is basically required high performance computing to build chemistry code as well as its simulation. HPC clustering is used to scale up the application performance and its availability. With the increasing demand of cluster computing many organizations are publishing innovative flavours in HPC cluster computing to boost up the application performance and scalability. High Performance Computing: which includes computers, networks, algorithms and environments to make such systems usable - range from small cluster of PCs to fastest supercomputer. WRF-CHEM model is successfully deployed on the HPC cluster over the years. As WRF-CHEM model dependency on the various clustering elements, like cps, memory, as well interconnected nodes.**

The purpose of this paper is to make benchmark for WRF-CHEM model on HPC cluster on the basis of different resolutions for specific regional data.

Keywords—WRF-CHEM; High Performance Computing (HPC); Portable Batch System (PBS); Benchmark

1. INTRODUCTION

Numerical weather prediction models are critical tools for forecasters. WRF is designed to provide real-time, extremely accurate and sophisticated weather analysis. WRF-CHEM is the chemistry model has been built to be consistent with the Weather Research Forecast model I/O Applications Program Interface. The Weather Research and Forecasting model (WRF hereafter) was developed at the National Center for Atmospheric Research (NCAR) that is operated by the University Corporation for Atmospheric Research (UCAR). WRF-CHEM model is used for Prediction and simulation of weather, or regional and local climate. Weather forecasts are routinely run on from hundreds to thousands of CPUs in parallel communicating over high-speed networks. WRF model is Weather Research Forecasting model, which is fully functioning modeling system for atmospheric Research & Operational research weather Prediction communities and WRF-CHEM has added chemistry predictions options which are mostly used by scientist of research industries. As WRF-CHEM requires more computations time it requires HPC computing. Clustering elements, such as the CPU, interconnects and the software libraries are crucial for enabling efficient predictions and high productivity. For efficient analysis, WRF requires high-performance computing system. Commodity clusters have become very important for high performance computing due to the price for performance, flexibility and scalability they can deliver.

Clustering is a group of loosely coupled commodity computers working together to achieve the same goal, maintain a Single System Image, good computational performance and Reliability. Clustering has been used in many fields including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics. Cluster analysis as such is not an automatic task, but an iterative process of knowledge discovery or interactive multi-objective optimization. It will often be necessary to modify preprocessing parameters until the result achieves the desired properties [1].

2. THE WRF-CHEM MODEL AND ITS ARCHITECTURE

2.1 The WRF-CHEM model system consists of following three major programs with are shown in Fig. 1:

- The WRF Pre-Processing System (WPS)
- WRF-Var data assimilation System
- WRF solver (ARW core only) including chemistry
- Post-Processing and Visualization tools

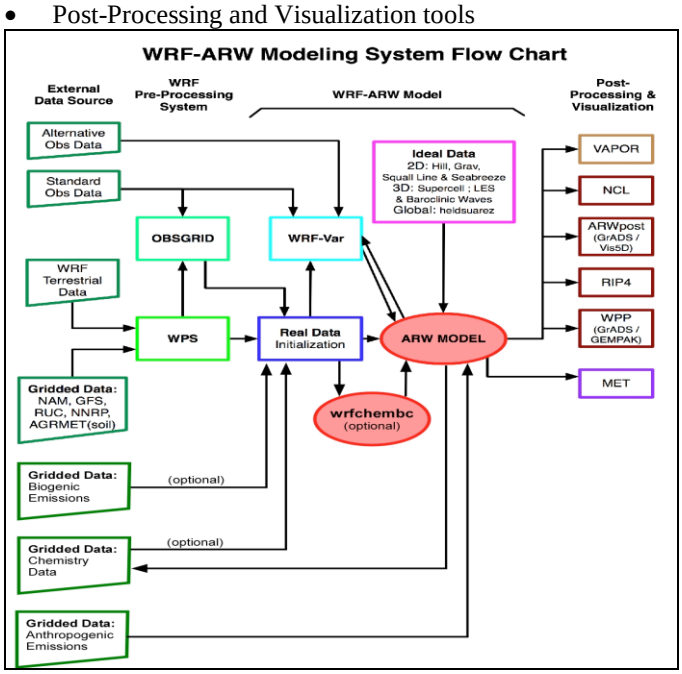

Fig. 1. WRF-CHEM modeling system architecture [2]

2.2 WRF-CHEM forecast should follow the below process,

1. Produce a WRF for particular domain with chemistry turned off.
2. Generate emissions data file for the domain and chemical mechanism of interest and write them into an intermediate file.
3. Convert emissions into a WRF chemical emissions input file.
4. Finally, produce WRF simulation for domain with chemistry.

For Advanced Research WRF with chemistry simulation requires a list of the supported combinations of hardware and software, required compilers, and scripting languages as well as postprocessing software. It cannot be guaranteed that chemistry will build successfully on all architectures that have been tested for the meteorological version of WRF [2].

2.3 The WRF-CHEM Model Difference

The difference with regular WRF comes from the chemistry part of the model needing to be provided additional gridded input data related to emissions. This additional input data is provided either by the WPS (dust emission fields), or read in during the real.exe initialization (e.g., biomass burning, biogenic emissions, GOCART background fields, etc.), or read in during the execution of the WRF solver (e.g., anthropogenic emissions, boundary conditions, volcanic emissions, etc.). And while some programs are provided in an attempt to aid the user in generation of these external input data files, as stated earlier, not all emissions choices are set-up to function for all possible namelist options related to the WRF-CHEM model [2].

WRF-CHEM forecast model is very complex model. Users of this model will need to change and set the environmental variables and available code as well as its standard configuration and compilations, depending on the various specifications of Linux systems to get it to function properly for their project.

WRF-CHEM model requires lots of memory computations and nodes for its additional variables settings. For that this model is deployed only in HPC computing environment. Also for its configuration and compilation process, distributed parallel memory option is selected to make it run.

3. PRL HPC Clustering Configration

Physical Research Laboratory (PRL) is a national Research institute for space and science, supported mainly by Department of Space, Government of India. PRL HPC Cluster Architecture: It is a 21 node cluster with 20 compute nodes and 1 master node with a peak performance of 3.2TF and a sustained performance of 2.2TF (approx.). It supports 64-bit Hardware and 32/64 bit software. It has different types of nodes like Backup Node, I/O Node, a Storage Node and a management node. The cluster is having a 10TB of usable storage based on FC disk drives (minimum 10k rpm). It has 20TB of raw storage with LTO Gen 4 Tape Library for Data Backup. The primary Network is Infiniband and the Secondary Network is Gigabit. There is as additional Management Switch for Node Management using intelligent platform Management Interface (IPMI). Fig. 2 shows PRL HPC cluster architecture [3].

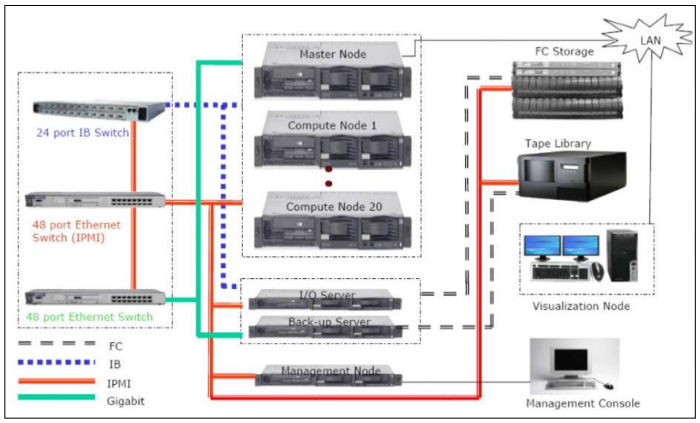

Fig. 2. PRL HPC Cluster architecture [3]

The Master node and all Compute nodes are installed with Red Hat Linux Enterprise Linux 5.1(2.6.18-53.e15) as operating System with Rock 5.1 as cluster management Tool.

- **Master Node:** It is having Quad core and Quad socket AMD Opteron 8360SE, 2.5GHz processor, 64 GB memory capacity and a capacity of 4*146GB SAS.
- **Computer Nodes:** All the 20 compute nodes are having Quad core and Quad socket AMD Opteron 8360SE, 2.5GHz processor, 64 GB memory capacity and Hard disc capacity of 2*73GB SAS.
- **Storage Node:** It is having Quad core Dual Socket Xeon E5420, 2.5 GHz processor, 8GB RAM, 4*72GB10K SAS.
- **Backup Node:** It is having Quad core Dual Socket Xenon E5420, 2.5 GHz processor, 8GB RAM, 2*120GB SATA.
- **Disk Array:** The EVA 4400 Disk Array is having 4 Disk Enclosures, 42*422GB 10K RPM FC Disks, total capacity of 16 TB (approx.) and a usable storage of 10TB.
- **IPMI Management Node:** Quad Core Dual Socket AMD Opteron2360, 2.5GHz processor, 4GB RAM, 2*160GB SATA.
- **Software used on PRL HPC Cluster:** Intel C,C++,Fortran, GNU C,C++,Fortran, Parallel Compiler Suites-INTEL MPI and OpenMP

Compilers, Intel Profiler and Analyzer and Intel Debugger Suite, Torque Scheduler Suite.

4. BENCHMARK CONFIGURATION OF PRL HPC FOR WRF-CHEM MODEL

We have configured and compiled WRF-CHEM model on PRL HPC cluster node successfully. For benchmarking of WRF-CHEM model on PRL HPC cluster we have taken real case study of Uttarakhand region for India.
We have made simulations for different resolutions on decided standard region.

*4.1 Following Table I. shows **the Benchmark configuration of PRL HPC Cluster** for WRF-CHEM:*

TABLE 1. BENCHMARK CONFIGRATION OF PRL HPC CLUSTER

Application	WRF-CHEM V3.5.1
Processors	Quad Core Quad socket AMD Opetron 8360 2.5GHZ Processor
Memory	64GB RAM per node
Hard Disk capacity	10TB HDD
OS	Red Hat Linux Enterprise 5.2
Compiler	Intel Ifort with icc (dmpar)
Message Passing Interface (MPI)	MPI 3.2.1 Library
Interconnect	Primary network is Infiniband and Secondary is Gigabit
Resource Management System Middleware (RMS)	Open Portable Batch System (OpenPBS)

*4.2 Following Table II. shows **Benchmark Time control and region specification** for WRF-CHEM Runs for Uttarakhand region of India:*

TABLE 2. BENCHMARK TIME CONTROL & REGION SPECIFICATION

Region:	**Start Date:**	**End Date:**
Uttarakhand	14th June, 2013	18th June, 2013

*4.3 The **Benchmark PBS specifications** of WRF-CHEM model for different resolutions run in MPI on PRL HPC:*

We have taken a standard PBS script for each run on HPC for Uttarakhand region. For that we have fixed the numbers of Nodes, cpus and memory for all runs which are show in Table III.

TABLE 3. BENCHMARK PBS SPECIFICATION

Nodes	3
CPUS	5
Memory	9GB
Total No. of CPUS for WRF task	5*3=15
Total No. of Memory for WRF task	9*3=27GB

4.4 Benchmark Result Comparisons of WRF-CHEM model:

Table IV. shows the different resolutions runs and their respected simulation times on standard PBS script specifications on PRL HPC.

TABLE 4. BENCHMARK RESULT COMPARISONS IN DIFFERENT RESOLUTIONS

Run 1: For 25km grid length in XY-directions, 61 end index in west-east directions, 61 end index in south-north directions. **Remark:** For 25km simulation, WRF-CHEM benchmark Time is **6hours** approx. on PRL HPC
Run 2: For 20km grid length in XY-directions, 75 end index in west-east directions, 75 end index in south-north directions. ***Remark:*** For 20km simulation, WRF-CHEM benchmark Time is **8hours** approx. on PRL HPC
Run 3: For 15km grid length in XY-directions, 100 end index in west-east directions, 100 end index in south-north directions. ***Remark:*** For 15km simulation, WRF-CHEM benchmark Time is **72hours** approx. on PRL HPC

5. SUMMARY AND CONCLUSION

We have discussed about the WRF-CHEM model architecture and how it is different from other forecasting models. WRF-CHEM model is used by the scientists in the institute for various weather researches and for scientific computations on the basis of change in the atmosphere. WRF-CHEM model simulations are very complex and time consuming. It is compulsory requirement for WRF-CHEM model to configure it in High performance cluster environment. After making successful establishment of WRF-CHEM model on PRL HPC we have concluded that WRF-CHEM simulation is very time consuming if we handle it in less numbers of cpus and memory. We also conclude that WRF-CHEM model can't run in serial environment as well as in shared memory environment. WRF-CHEM simulation runs in distributed parallel environment only on HPC cluster and giving us better performance with chemistry data.

We have also compared our benchmark results with the same HPC hardware configuration but with the different Kilometers resolutions and numbers of points of grid. So, from these experiments we are now finalizing that with increasing of resolutions WRF-CHEM model decreases its performance, so to increase simulation performance we need more HPC nodes

as well as high memory space & computing power. If we can improve the nodes with sufficient memory and computing power of cpus then we can achieve best & fast simulation results.

Acknowledgments

A special thanks to Dr. Prashant Kumar, who is the Scientist at Atmospheric Science Division at Space Applications Centre in Indian Space Research Organization, Ahmedabad for providing his guidance to establish WRF-CHEM model at PRL.

REFERENCES

[1] Foram Shukla, Prof. Gayatri Jain, and Dr. Varun Sheel, "Trends of optimizing the usage of parallel HPC cluster in various domains," in ijarcsse, vol. III,issue 10,pp. 623-628,October 2013.

[2] Dr. Steven E. Peckham et al., "WRF/Chem version 3.5 User guide",July 2013.

[3] PRL HPC computing: "http://www.prl.res.in/",September 2013 .

[4] High performance computing: "http://www.cdac.in/".

[5] El Amrani, Chaker, and Ian M. Hedgecock. "Implementation of the WRF-Chem model in Grid Computing and GPU for regional air quality forecasting."*Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*. IEEE, 2012.

[6] Shainer, Gilad, et al. "Weather Research and Forecast (WRF) Model Performance and Profiling Analysis on Advanced Multi-core HPC Clusters."*10th LCI ICHPCC* (2009).

[7] Yeo, Chee Shin, et al. "Cluster Computing: High-Performance, High-Availability, and High-Throughput Processing on a Network of Computers." *Handbook of nature-inspired and innovative computing*. Springer US, 2006. 521-551.