Coupling NASA Advanced Multi-Scale Modeling and Concurrent Visualization Systems for Improving Predictions of Tropical High-Impact Weather (CAMVis) AIST08-0049

B.-W. Shen[1,2], B. Green[3], W.-K. Tao[2], C. Henze[3], S. Cheung[3], J.-L. Li[4], P. Mehrotra[3]
[1]UMD/ESSIC, [2]NASA/GSFC, [3]NASA/ARC, [4]NASA/JPL

Abstract:

The ESTO AIST project, *Coupling NASA Advanced Multi-Scale Modeling and Concurrent Visualization Systems for Improving Predictions of Tropical High-Impact Weather (CAMVis)*, has completed its first phase this year. The project seeks to improve high-impact tropical weather prediction by seamlessly integrating NASA technologies (e.g., multiscale modeling, supercomputing and visualization) to create a real-time, high-resolution, weather prediction tool. CAMVis will be able to produce high-resolution simulations, inter-compare the simulations with multi-satellite observations (e.g., QuikSCAT winds and TRMM precipitation), and to provide advanced 3D visualizations of atmospheric moist thermodynamic processes and cloud-radiation-aerosol interactions.

"*Extreme Event Warning*", among scenarios in Decadal Survey Missions report (NRC, 2007), was identified as one of top priority scenarios by the advanced data progressing group (ADP) at the ESTO AIST PI Workshop 2010. This scenario focuses on "*discovering predictive relationships between meteorological and climatologically events and less obvious precursor conditions from massive data set.*" To achieve this scientific goal, our approach is to build the the CAMVis information system by (1) integrating the existing NASA technologies such as NASA multi- scale model system, Goddard Cloud Ensemble model (GCE), the finite-volume General Circulation Model (fvGCM), and the concurrent visualization (CV) systems; (2) improving parallel scalability of the coupled multi-scale modeling system to take full advantage of the next-generation peta-scale supercomputers (e.g., Pleiades supercomputer); (3) significantly streamlining data flow for fast processing and 3D visualizations; (4) developing visualization modules for fusion of NASA TRMM precipitation and QuikSCAT winds. During the first year of the AIST Project CAMVis, we have finished the following tasks: (i) initial coupling of the high-

resolution fvGCM and CV systems on the new NASA Pleiades supercomputer; (ii) development of the meta-global Goddard Ensemble Model (mg-GCE); (3) development of CV system version 2.0 and implementation of parallel "M-on-N" data transfer between "M" computing nodes and "N" visualization nodes.

In this paper, we will first discuss the current progress of the aforementioned components. Secondly, with the initial coupled system (fvGCM-CV), we will show that the CAMVis information system has been used to produce 3D visualizations for providing the insightful understanding of multiscale processes and their multiscale interactions, which are key to realistically predicting the formation of tropical cyclones (TCs) such as twin TC and Nargis (2008) and large-scale tropical weather system such as Madden-Julian Oscillations. TC Nargis is one of the 10 deadliest tropical cyclones in recorded history, causing 133,000 fatalities and $10 billion dollars in damage in Burma (Myanmar) in early May 2008. Finally, we will discuss future tasks that would enhance and enable CAMVis to support the following NRC Decadal Survey Earth Science missions, including CLARREO, ACE, PATH, 3D-Winds.

1. Introduction

To achieve the goal of *"Extreme Event Warning"*, we start with making an attempt of extending the lead time and reliability of hurricane forecasts, which is important for saving lives and mitigating economic damage. The urgent need for doing this is evidenced by extreme weather events such as the 2005 Hurricane Katrina (Shen et al., 2006a) and 2008 Tropical Cyclones Nargis (Shen et al, 2010a), which caused tremendous damage and numerous fatalities. It has been suggested that large-scale tropical weather systems such as Madden-Julian Oscillations (MJOs, Madden and Julian, 1973), monsoonal circulation and tropical easterly waves may regulate the activities of tropical cyclones (TCs[1], e.g., Maloney and Hartmann, 2000; Shen et al., 2010a). To this end, improving predictions of these large-scale flows might be helpful for extending the lead-time of TC prediction. However, due to limited computing resources, it has been very challenging to accurately improve these tropical weather systems with traditional global models. Two of the

[1] Depending on their location, TCs are referred to by other names, such as hurricanes (in the Atlantic region), typhoons (in the West Pacific region), tropical storms, cyclonic storms, and tropical depressions.

major limiting factors in these models are insufficient grid spacing and physics parameterizations (e.g., cumulus parameterizations, CP[2]).

Thanks to recent advances in global weather/climate modeling and supercomputing, there is now great potential for mitigating the aforementioned issues (e.g., Kerr 2006). In late 2004, NASA's Columbia supercomputer started operation (Biswas et al., 2007), providing groundbreaking computing power for Earth modeling. Later, the NASA high-end concurrent visualization (CV) system version 1 (Ellsworth et al., 2006), in which model outputs are extracted for analysis while the simulation is still running, was developed as a powerful tool for efficiently processing and visualizing massive volumes of high spatial- and temporal-resolution model outputs. In late 2008, a new supercomputer, Pleiades, was installed at NASA Ames Research Center, and provides 10 times Columbia's computing power. Enabled by these advanced computational technologies, a high-resolution (~10km) global model (the finite-volume general circulation model, known as fvGCM) was deployed and used to generate remarkable forecasts of intense hurricanes (Atlas et al., 2005; Shen et al., 2006a,b,c; Shen et al, 2009a, Shen et al., 2010a). More importantly, an innovative approach that applies a massive number of cloud-resolving models (CRMs, the Goddard Cloud Ensemble model, Tao and Simpson, 1993; Tao et al., 1993) and the fvGCM has been proposed and used to overcome the cumulus parameterization (CP) deadlock[3] in GCMs (e.g., Tao et al., 2008, 2009b). The second approach is called the multiscale modeling framework (MMF) or super-parameterization, wherein a CRM is run (at a resolution of 4km, currently) in place of the CP in a GCM grid (Randall et al., 2003). As a result, the MMF has the combined advantages of the global coverage of a GCM and the sophisticated microphysical processes of a CRM. Both the high-resolution global model and MMF can be run with no dependence on CPs, and are used to examine the impacts of resolved convections (over ocean and land) and their interaction with radiation on the simulations of the aforementioned large-scale tropical systems.

In section 2 of this article, we introduce the supercomputing, concurrent visualization and global modeling technologies at NASA. We then discuss in section 3 how the CAMVis system

[2] Typical resolutions used in these GCMs were too coarse to explicitly resolve small-scale moist thermodynamic processes associated with cloud motions, and therefore "cumulus parameterizations" (CPs) were often designed to emulate the statistical effects of unresolved subgrid-scale cloud motion.

[3] It has been shown that the development of CPs has been very slow, and their performance is a major limiting factor in weather/hurricane simulations.

built with these technologies can help improve short-term (5~10-day) forecasts of TCs and extended-range (30-day) simulations of large-scale tropical weather systems such as MJOs, and provide insightful understanding of multiple physical processes and their multi-scale interactions. We conclude with a summary and discussion of future plans.

2. Supercomputing and Modeling Technology at NASA

2.1 Supercomputers

Figure 1: NASA Supercomputers. Left panel: Columbia supercomputer with (in late 2004): (1) 20 SGI Altix superclusters, each with 512 CPUs; (2) 10,240 Intel Itanium II CPUs; (3) 20 TB total memory (e.g., Biswas et al., 2007). Right panel: Pleiades Supercomputer (ranked 3rd in late 2008) consists of 51,200 cores (with quad-core Xeon 5472 processors) in total, 50+ TB memory, and 500+ TB disk spaces.

In late 2004, the Columbia supercomputer (Biswas et al., 2007) came into operation with a theoretical peak performance of 60 Tflop/s (trillion floating-point operations per second) at NASA Ames Research Center (ARC). It consists of twenty 512-CPU nodes, which give 10,240 CPUs and 20 terabytes (TB) of memory. Columbia achieved a performance of 51.9 Tflop/s with the LINPACK (Linear Algebra PACKage) benchmark and was ranked second on the TOP500 list in late 2004 (Figure 1). These large-scale computing capabilities enable complex problems to be resolved with large-scale modeling systems (e.g., Tao et al. 2008; Shen et al. 2006a). In late 2008, a 51,200-core SGI Altix ICE system, Pleiades (609 Tflop/s, peak), was built as one of the most

powerful general-purpose supercomputers. This newly built system, which provides 10 times the computing power of Columbia, is expected to speed up science discovery at an unprecedented pace.

2.2 Concurrent Visualization System

It is well known that a substantial increase in data volume produced by high-resolution Earth modeling systems poses a great challenge to stage, handle, and manage these model outputs and compare them with satellite data. Our extensive experience with large-scale modeling systems indicates that speeding up pre- and post-processing for model's inputs and outputs is as important as scaling the model's performance. We believe that to efficiently handle these massive datasets, from terabytes for short-term runs to petabytes for long-term runs, requires an innovative thought-process and approach. A technique that could achieve this goal, and has previously met with great success in visualizing high-volume data from very high-resolution simulations, is concurrent visualization (CV; Ellsworth et al., 2006; Green et al., 2010). In CV, a simulation code is instrumented such that its data can be extracted for analysis while the simulation is running, without having to write the data to disk. By avoiding filesystem I/O and storage costs, CV has the benefit of providing much higher temporal resolution than is possible with traditional post-processing, enabling every timestep of a very high-resolution simulation to be captured for analysis, when only a small fraction of the timesteps could reasonably be saved to disk. The other main benefit of CV is that it provides a view of a simulation in progress, which may be useful for application monitoring or steering. This can help prevent serious job failures that might otherwise cause waste of system resources.

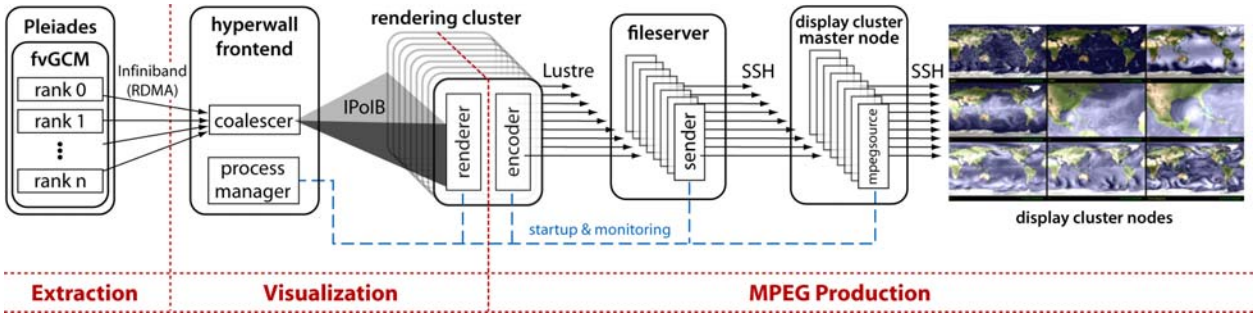


Figure 2: The newly developed Concurrent Visualization (CV) System (Version 2). Rounded rectangles indicate systems, and rectangles indicate processes. The whole system (from left to

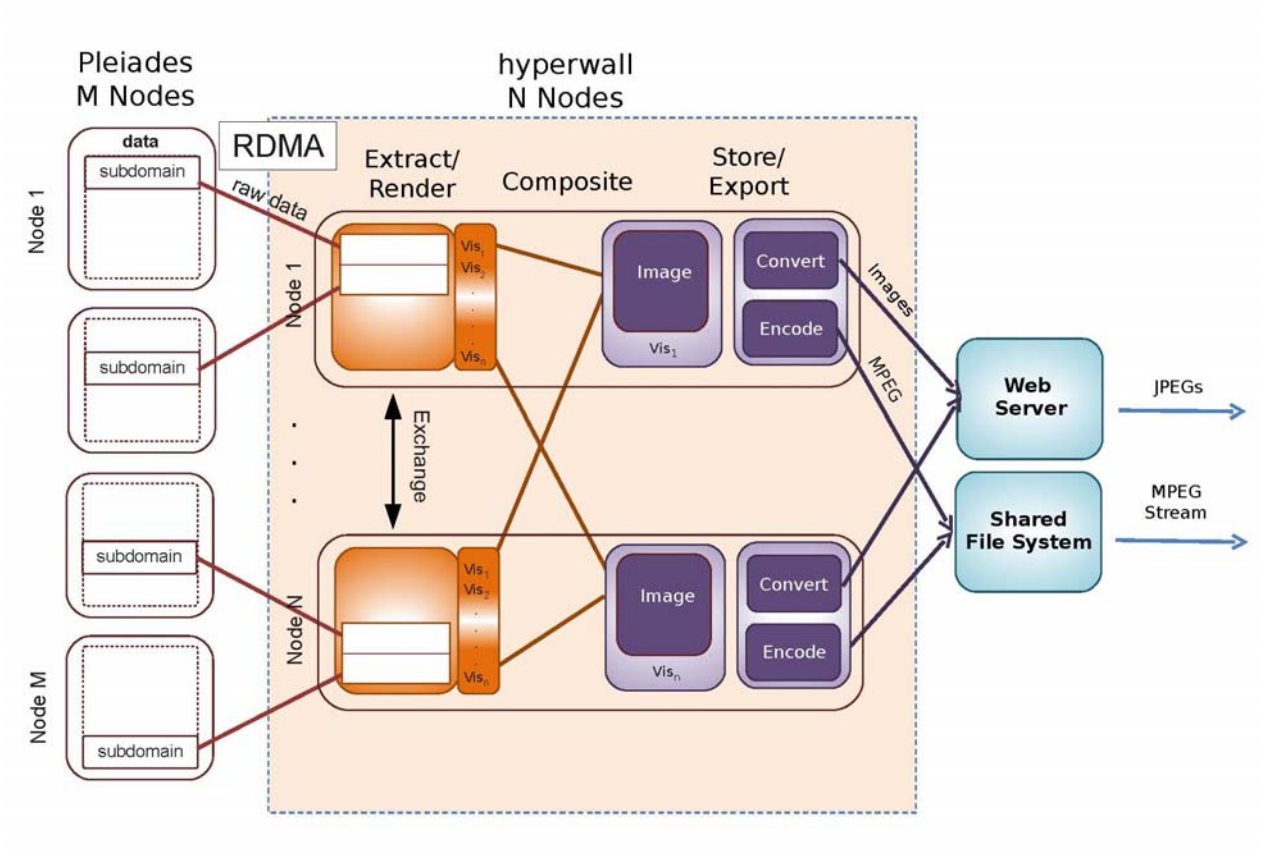
right panels) consists of a computing node ("Pleiades Node"), a middle-layer system ("coalescer"), and the hyperwall-2 128-node 8-core/node rendering cluster. These systems are used for data extraction, handling, and visualization; and for MPEG image production and visualization display. (Green et al., 2010).

As part of the Columbia project, during 2005, the CV technology was first developed and integrated into the high-resolution fvGCM on the original hyperwall system (49 screens), where at least 9 visualization clients were running. A 3x3-screen "mini-hyperwall" was used for looping the resulting movies. Recently, an new improved CV system version 2 has been deployed (Figure 2), which consists of a front-end system for data extraction ("coalescer"), a middle-layer system for data handling and data rendering, and a back-end system for data display. "Extractions" include domain slices or subvolumes, cutting planes, isosurfaces, streamlines, and other feature-extraction products. The size of an extract can vary widely, depending on what features are captured, while the size of a movie is determined only by the image resolution and the compression level achieved during encoding. The great advantage of extracts is that they represent an intermediate data product, which can be loaded into a viewer at a later time for interactive analysis.

As of June 2008, NASA's 128-screen hyperwall-2, capable of rendering one-quarter-billion pixel graphics, was built at NASA ARC as one of the world's highest resolution scientific visualization and data exploration systems. Compared to the original 49 screens and 100BASE-T interconnect, the hyperwall-2 has 128 screens with modern graphics cards, an InfiniBand interconnect, and is fully integrated into the NAS supercomputing environment. The hyperwall-2's 1,024-CPU cores and 475 terabytes of fast disk provide an excellent environment for parallel feature extraction and extract storage. In addition to greatly enhanced visualization power, hyperwall-2's high-speed interconnect makes fully 3D concurrent visualization possible.

To efficiently exchange data between the computing and visualization nodes, we have implemented the M-on-N configuration for the concurrent visualization pipeline, as shown in Figure 3. The boxes on the left represent the multiple (M) MPI processes of a fvGCM job, each responsible for simulation within a sub-domain (a portion of the whole domain). At startup, each MPI process in the fvGCM job creates a connection with the Infiniband RDMA (remote data memory access) protocol to one of the N MPI processes spawned from the hyperwall job, using an M-on-N mapping where M>=N. At the end of each timestep, raw output on a computing node

is transferred directly via its RDMA connection to its corresponding peer process on the specific hyperwall node. The hyperwall job then performs feature extraction and "sort-last" rendering in parallel: each of its child MPI processes renders an image from its own data, then these "individual" images from all of the child MPI processes are sorted and composited into a complete image in a PPM format, which can be passed to an encoder for movie generation. Finally, a series of these complete images are converted to JPEG, which can be easily delivered to a web server for display.



<u>Figure 3:</u> The M-on-N configuration for parallel data communications between the M computing nodes and N visualization nodes in the concurrent visualization (CV) system version 2 (Green et al., 2010).

In order to maximize the results from a single simulation run, multiple products are usually generated, representing various fields and regions of interest, numerous feature-extraction and visualization techniques. As part of the CV pipeline, resulting animations are streamed, as they are being generated, to the remote displays at the facilities of the principal scientists. When time-stamped outputs arrive, the visualizations are processed in round-robin fashion (Vis1, Vis2, ...

Visn), producing one image per visualization request. The destination node for the final composite image is also assigned in a round-robin fashion, so that the encoders are spread out across the cluster.

Depending on the visualization produced, some data exchange may occur within the hyperwall nodes. Visualizations such as scalar volume rendering, cutting planes, and iso-surfaces are easily implemented within the "sort-last" renderer, with only ghost-cell exchange needed for the sub-domain boundaries. However, for performing vector visualization techniques within the subregions, which may cover several sub-domains on different nodes, it may be convenient to fully reconstruct a subregion of interest on each of the visualization nodes.

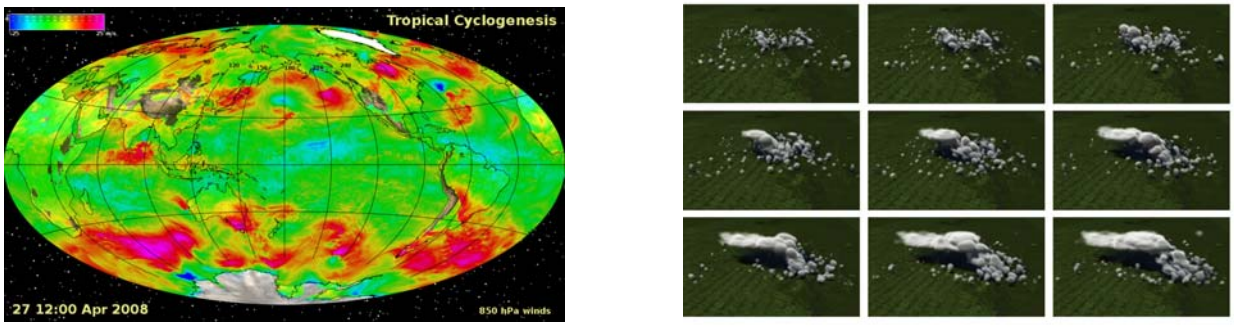
2.3 Multiscale Modeling System

The multi-scale modeling system with unified physics (Tao et al., 2008) has been successfully developed at NASA Goddard Space Flight Center (GSFC) and deployed on the Columbia and Pleiades supercomputers. As shown in Figure 4, the system consists of the global model (finite-volume GCM, fvGCM; Lin and Shen et al., 2003; Lin 2004; Atlas et al., 2005; Shen et al., 2006a) and thousands of copies of a cloud model (Goddard Cumulus Ensemble model, GCE; e.g., Tao and Simpson, 1993; Tao et al., 1993). With the current model configurations, 13,104 GCEs are running concurrently to explicitly simulate cloud processes in the global environment, providing cloud feedbacks to the atmospheric state in the fvGCM. Recently, a satellite simulator has been integrated into the system (Matsui et al., 2008), which provides a powerful tool for comparing model results with satellite data.

The high-resolution fvGCM was first deployed on the Columbia supercomputer (and later on Pleiades), producing remarkable forecasts of intense hurricanes in 2004 and 2005 (e.g., Atlas et al., 2005; Shen et al., 2006a,b,c; Shen et al., 2009a; Shen et al., 2010a). The parallelization of the fvGCM was carefully designed to achieve efficiency, scalability, flexibility, and portability. Its implementation had a distributed- and shared-memory two-level parallelism, including a coarse-grained parallelism with MPI and fine-grained parallelism with OpenMP (Putman et al., 2005). One of the prominent features in the implementation is to allow multi-threaded data communication.

Both of the MMF and high-resolution fvGCM, which can be run with no dependence on CPs, are powerful tools to examine and understand the impacts of resolved convections (over ocean

and land) and their interaction with radiation on the simulations of high-impact tropical weather such as MJOs and TCs, aimed at improving predictions of these weather system.



<u>Figure 4:</u> The NASA global multiscale modeling system, consisting of the global model (fvGCM) and thousands of copies of cloud models (GCE). Left panel: visualization of a 5-day low-level wind simulation with the global model, showing the formation of severe tropical storm Nargis (2008) in the Indian Ocean. Right panel: 3D visualization of cloud processes with 9 cloud models in a $6^o x 7.5^o$ area, giving 13,104 cloud models in the global environment.

2.3.1 Computational enhancement:

At the runtime of MMF's execution, 95% or more of the total wall-time for running the MMF is spent on the multiple copies of the GCEs. Thus, wall-time could be significantly reduced by efficiently distributing the large number of GCEs over a massive number of processors on a supercomputer. However, the original implementation, in which each of 13104 GCEs is embedded on a grid point in the fvGCM, has very limited parallel scalability with a total number of CPUs up to 30. To overcome this difficulty, a different strategic approach is proposed to couple the fvGCM and GCEs as discussed below. From a computational perspective, the concept of "embedded GCEs" should be completely forgotten, as it restricts the view on the data parallelism of the fvGCM. Instead, the 13,104 GCEs should be viewed as a *meta global GCE* (mgGCE) in a *meta gridpoint system*. With this concept in mind, each of the two distinct parts (or "components"; the fvGCM and mgGCE) in the MMF could have its own scaling properties. Since most of wall-time in MMF runs was spent on the GCEs, a scalable mgGCE could substantially reduce the wall-time. In addition, it becomes feasible to implement an MPMD (multiple programs multiple data) parallelism for the MMF with the mgGCE and the fvGCM. The technical

approaches for this implementation are briefly summarized as follows: (1) a master process allocates a shared memory arena for data redistribution between the fvGCM and mgGCE by calling the Unix *mmap* function; (2) the master process spawns multiple (parent) processes with a 1D domain decomposition in the y direction by a series of Unix *fork* system calls; (3) each of these parent processes then forks several child processes with another 1D domain decomposition along the x direction; (4) data gathering in the mgGCE is done along the x direction and then the y direction; (5) synchronization is implemented with the atomic *__sync_add_and_fetch* function call on the Columbia supercomputer. While steps (1), (2), and (5) were previously used in MLP (multiple level parallelism) by Taft (2001), this methodology is now extended to the multi-component system (see also Shen et al., 2009c).

Figure 5 shows preliminary benchmarks with very promising scalability up to 364 CPUs. Here the speedup is determined by T30/T, where T is the wall time to perform a 5-day forecast with the MMF and T30 the time spent using 30 CPUs. The run with 30 CPUs was chosen as a baseline simply because this configuration was previously used for production runs. A speedup of (3.93, 7.28, and 12.43) is obtained by increasing the number of CPUs from 30 to (91, 182, and 364) CPUs, respectively.

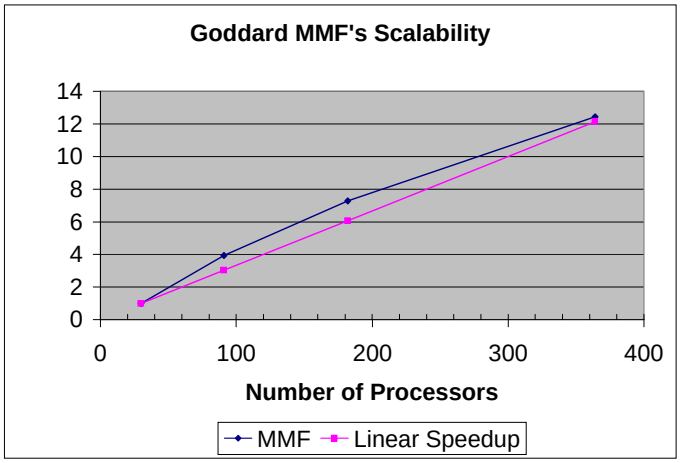


Figure 5: Scalability of the Goddard MMF with a proof-of-concept parallel implementation. This figure shows that a linear speedup is obtained as the number of CPUs increases from 30 to 364. The original MMF could use only 30 CPUs production runs (see also Shen et al., 2009c).

3. Scientific Applications:

In the following, numerical results and visualizations from the CAMVis system with improved convective/cloud processes are discussed to illustrate (1) improved short-term (5~7 days) predictions for the formation of twin TCs associated with a large-scale MJO; (2) insightful

visualizations of multiple processes and their scale interactions that lead to the formation of very severe tropical storm Nargis (2008); (3) improved extended-range (~30 days) simulations of a large-scale MJO; and (4) inter-comparisons between model simulations and Satellite measurements at comparable resolutions.

3.1 Simulations of Twin Tropical Cyclone Formation associated with an MJO

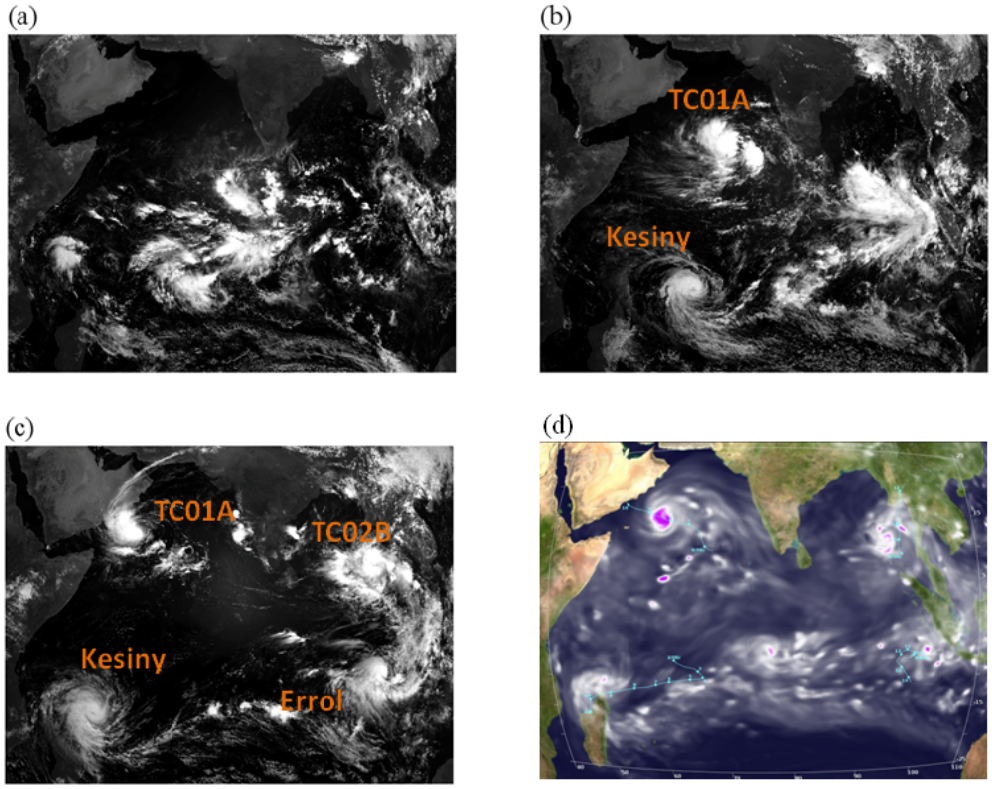


Figure 6: Predictions regarding the formation of twin tropical cyclones in the Indian Ocean: (a) MJO-organized convection over the Indian Ocean at 0630 UTC 1 May 2002. When the MJO moved eastward, two pairs of twin TCs appeared sequentially on 6 May (b) and 9 May (c), including TC 01A, Kesiny, TC 02B and Errol. Two TCs (01A and 02B) with anti-clockwise circulation appeared in the Northern Hemisphere, while two TCs (Kesiny and Errol) with clockwise circulation in the Southern Hemisphere; (d) Four-day forecasts of total precipitable water, showing realistic simulations of TC's formation and movement (see Shen et al., 2010b for details).

It has been documented that the nearly simultaneous formation of two TCs straddling the equator at low latitudes occasionally may occur in the Indian Ocean and West Pacific Ocean (e.g.,

Lander 1990). These TCs are called “twins” as they are nearly symmetric with respect to the equator. Previous studies showed that this twin TC activity can be modulated by the eastward propagation of the Madden-Julian Oscillation (MJO, e.g., Maloney and Hartmann 2000). For example, in early May 2002, large-scale organized convection associated with an MJO event was observed in the Indian Ocean (Figure 6a). While the MJO was continuously progressing eastward, two pairs of twin TCs (Figures 6b-6c) appeared. To capture the genesis, a 10-day forecast was initialized at 0000 UTC 6 May, as shown in Figure 6d. It is found that the genesis and movement of three of these TCs (02B, 01A and Kesiny) are simulated realistically; however, for the southern entity of the second pair of twin TCs (Errol), only less-organized convection is simulated.

3.2: 3D visualization of TC Nargis (2008)

Very severe cyclonic storm Nargis (2008) is the deadliest named tropical cyclone (TC) in the North Indian Ocean Basin, causing over 133,000 fatalities and $10 billion in damage. An increased lead time in the prediction of TC Nargis would have increased the warning time and may therefore have saved lives and reduced economic damage. Our global high-resolution simulations using real data show that the initial formation and intensity variations of TC Nargis can be realistically predicted at a lead time of up to five days. Experiments also suggest that the accurate representations of environmental flows such as 1) a westerly wind burst associated with a Madden-Julian Oscillation (MJO) is important for predicting the formation of this kind of TC. Other favorable factors for the formation and intensification of TC Nargis include 2) an enhanced monsoon circulation with a zero wind shear line, 3) good upper-level outflow with anti-cyclonic wind shear between 200 and 850 hPa, and 4) low-level moisture convergence. More detailed discussions can be found in Shen et al., (2010a). To provide a simplified view of these multiple processes, 3D visualizations could prove very powerful.

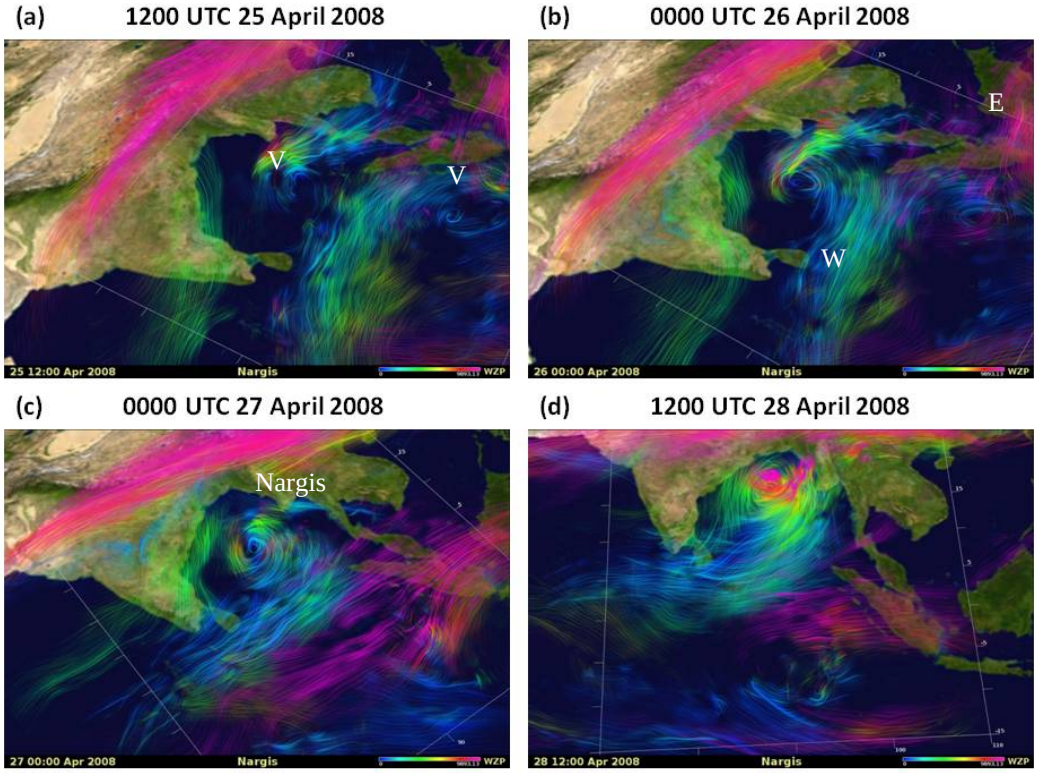


Figure 7: Realistic 7-day simulations of the formation and initial intensification of TC Nargis (2008) initialized at 0000 UTC 22 April 2008, showing streamlines at different levels. Low-level winds are in blue and upper-level winds in red: (a) formation of a pair of low-level mesoscale vortices (labeled in 'V') at 84h simulation. (b) intensification of the northern vortex (to the left) (c); formation of TC Nargis associated with the enhancement of the northern vortex; (d) intensification of TC Nargis associated with upper-level outflow and moist processes, indicated by the enhanced upper-level outflow circulation. Approaching easterly upper-level winds (labeled in 'E') increase the vertical wind shear, suppressing the enhancement of the southern vortex (to the right). (see Shen et al., 2010a for details).

In contrast to the twin TC case, simulating and understanding processes for developing TC Nargis and the non-developing vortex (the counterpart of Nargis in the Southern Hemisphere) are equally important, from a numerical prediction perspective. As the fvGCM was instrumented for the CV version 2 (Green et. al., 2010), and hyperwall, a set of 3D, high-temporal-resolution animations were produced for illustrating the above. Snapshots of streamline visualization at different vertical levels are shown in Figure 7. Low-level winds are shown in blue, and upper-

level winds in red. In Figure 7a, ending at 1200 UTC 25 April 25 2002, a pair of low-level vortices (labeled in 'V') appeared in the Northern and Southern Hemispheres, showing the potential for the formation of a pair of twin TCs. As time proceeded, the (low-level) westerly wind belt/burst (labeled in 'W') moved northward, enhancing the horizontal wind shear and therefore intensifying the northern vortex into the TC Nargis (Figure 7b-c). With other favorable factors, including good upper-level outflow, the TC Nargis continued to intensify (Figure 7d). In contrast, at 0000 UTC April 26 (Figure 7b), an upper-level easterly wind (labeled in 'E'), which moved on the top of the southern vortex, increased vertical wind shear and therefore suppressed the enhancement of the southern vortex (Figure 7b). With other unfavorable factors (such as closeness to the Equator), no TC formed in the Southern Hemisphere during the lifetime of TC Nargis.

3.3 MJO Simulations with the MMF

In the previous subsection, we have shown that the modulation of TC activities by MJOs can be realistically simulated with our modeling system. Here, we discuss the model's performance in simulating an MJO. It is known that accurate prediction of tropical activity at sub-seasonal scales (~30 days) is a crucial goal for extending numerical weather prediction (NWP) beyond two weeks. Among the challenges of this goal is accurate forecasting of an MJO. With a 45- to 60-day time scale, eastward-propagating MJOs, which are typically characterized by deep convection originating over the Indian Ocean, have one of the most prominent large-scale features of the tropical general circulation. Current understanding indicates that multiple processes, including moisture convergence, surface heat and moisture fluxes, cloud radiation feedback, and convection-water vapor feedback, are all important for the MJO's initiation, intensification, and propagation. The new multi-scale modeling framework (MMF) provides an innovative approach to investigating these multiple processes and multi-scale interactions. Figure 8 shows the 30-day simulation of an MJO event initialized at 0000 UTC December 13 2006, illustrating that the life cycle of the MJO is successfully captured, and suggesting that the model has a potential for examining the impact of an MJO on climate simulations.

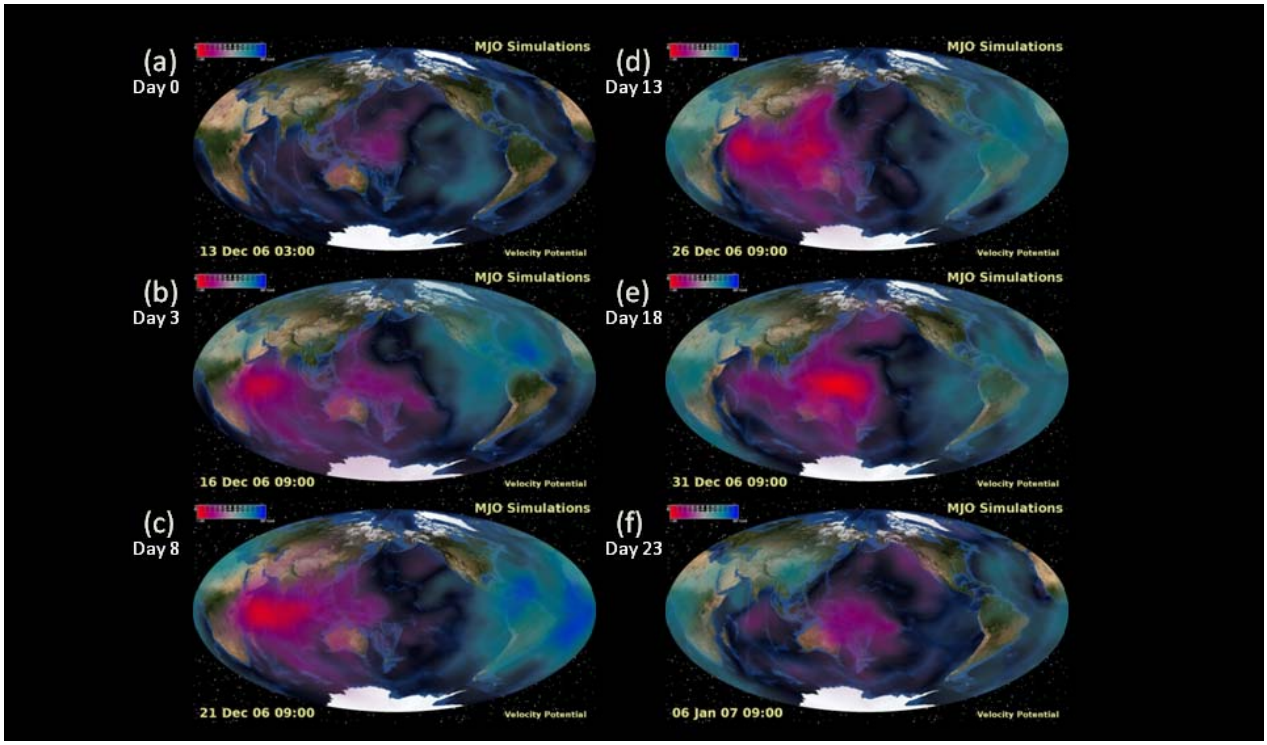


Figure 8: A 30-day simulation of an MJO initialized at 0000 UTC December 13, 2006, as shown in 200 hpa velocity potential. This simulation captures several major features usually associated with an MJO: (1) initiation of large-scale organized convection in the Indian Ocean in panel (b), (2) intensification as shown in panel (c), (3) slow propagation (prior to reaching the Maritime continent), (4) followed by fast propagation, and (5) weakening. However, this simulated MJO also produces stronger vertical motion than does the NCEP/GSF reanalysis.

3.4: Inter-comparisons between model simulations and Satellite measurements

As precipitations are a good indicator for energy source of intensifying TC and low-level wind speeds for the measurement of TC intensity, data fusion of NASA TRMM precipitations and QuikSCAT winds into the CAMVis system becomes valuable for inter-comparisons of high-resolution model simulations with satellite measurements. During the first year execution of the project, we have developed data conversion and visualization modules for this purpose. Figure 9a and 9b show the TRMM precipitation and QuikSCAT winds during the lifetime of TC Nargis (2008), respectively.

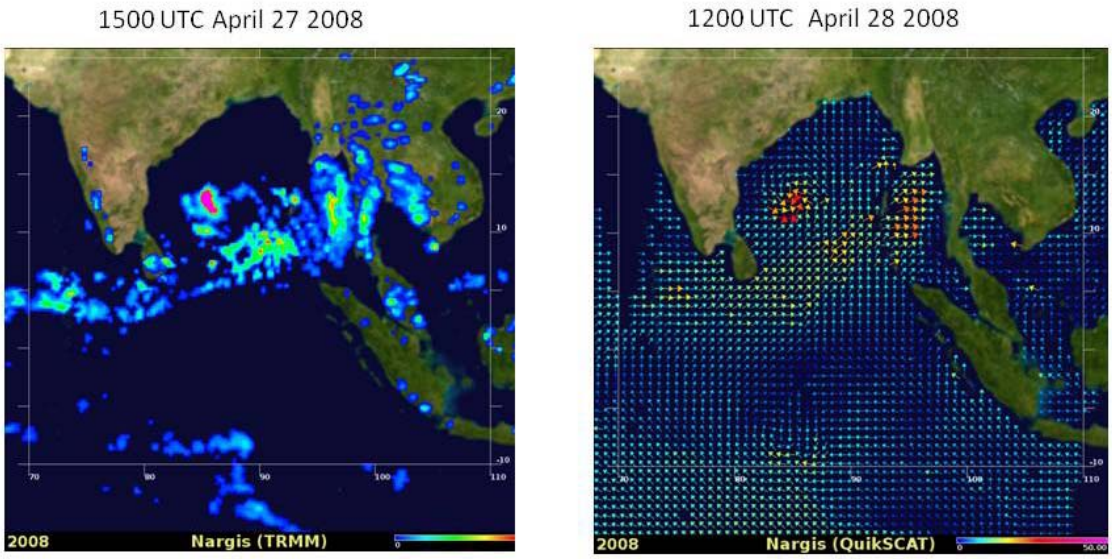
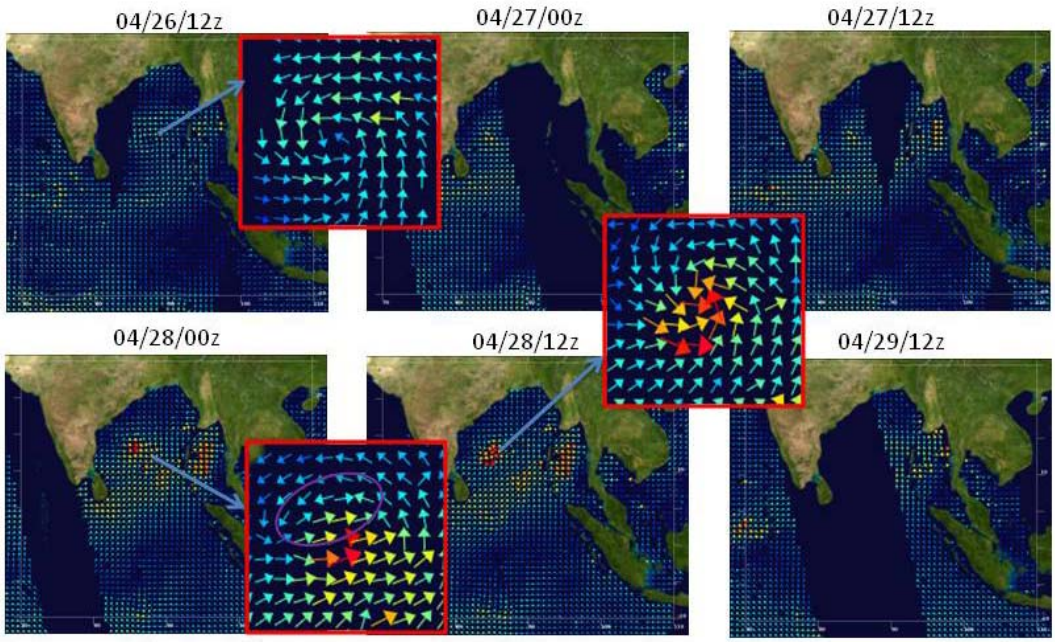


Figure 9: *Initial implementation of a visualization module into the CAMVis information system, including data convert and vector plotter for TRMM satellite-derived precipitation (left panel) and QuikSCAT winds (right panel), respectively. These figures show the TC Nargis (2008).*

With the new capability for data fusion, QuikSCAT winds for Nargis (2008) are inter-compared with high-resolution model simulations, aimed at understanding the data consistent accuracy in the representation of mesoscale vortex circulation and thus improving formation prediction. The assurance of data continuity (or consistency) is important for accurately tracing a TC movement or identifying its formation. From the zoom-in panels of Figure 10, it is clearly shown that the changes of vortex structure are not smooth (e.g., less realistic vortex in Figure 10d), suggesting the potential rainfall contamination on the derived wind distributions. This might impact the detection on the formation of a TC.



<u>Figure 10:</u> Vector visualizations of NASA QuikSCAT winds during the initial formation and intensification of Nargis (2008) from 1200 UTC April 26, 2009 at a time interval of 12 hours. Zoom-in windows are used to track the evolution of the mesoscale vortex with a closed circulation at 04/26/12z, 04/28/00z and 04/28/12z, respectively.

4. Concluding Remarks:

To support NASA missions and reduce the time to scientific discovery, we propose to seamlessly integrate the NASA advanced modeling and supercomputing technologies. Our plan is to improve the system's performance by taking full advantage of Pleiades' computing power, and to improve the simulations of cloud processes with 3D cloud models (GCEs) and to implement and test more sophisticated cloud (microphysical) schemes. The coupled system will be improved to address the interactions of cloud, radiation, and aerosol, in order to advance our understanding of the detailed 3D structure of these fields, and to investigate their impacts on tropical weather predictions by inter-comparing these high-resolution simulations with NASA high-resolution satellite observations. These satellites include current missions such as Tropical Rainfall Measuring Mission (TRMM), and Quick Scatterometer (QuikSCAT), and future missions such as Global Precipitation Measurement (GPM) and Aerosol-Cloud Ecosystems

(ACE), and 3D-Winds missions described in the NRC Decadal Survey (2007). In this paper, we give a summary on the project progresses for the development and deployment of the CAMVis and on its application to improving simulations of high-impact tropical weather systems including mesoscale tropical cyclones and large-scale Madden-Julian Oscillations (MJOs).

Previously, each of the individual components of the CAMVis system has been successfully developed and tested since the Columbia supercomputer came into operation in late 2004. During the first year of execution, starting from March 2009, initial deployment of the coupled system on the new NASA supercomputer Pleiades has been completed. As the mutiple-scale modeling systems can simulate weather and climate at high spatial and temporal resolutions, coupling these modeling and concurrent visualization systems can help process massive volumes of output efficiently (with no need of writing intermediate data into disks) and provide insightful understanding of the complicated physical processes. The CV system is equipped with the 128-screen hyperwall-2, and is connected via high-speed InfiniBand to the Pleiades supercomputers. The CV system has the following primary benefits: The first enables one to monitor system runtime status and thereby prevent serious failures that could cause waste of system resources. The second is to enable the use of much higher temporal resolution, as I/O and storage space requirements are largely obviated. Thirdly, it can help visualize the complicated physical processes, as shown with 3D visualizations for the formation of Nargis (2008).

For scientific applications with the CAMVis system, we first illustrate that the modeling systems can help improve the prediction of tropical cyclogenesis by improving simulations of multiple physical processes at multiple temporal and spatial scales (such as an MJO) and their multi-scale interactions. We then show that improved representations of multi-scale interactions are the key to extending the lead time of predicting severe cyclonic storm Nargis (2008), which devastated Burma (Myanmar) in May 2008 and caused massive damages and fatalities. In addition, we demonstrate that with a 30-day simulation, the modeling system is capable of simulating the life cycle of the MJO event in 2006, aimed at extending the lead time of short-term TC prediction. The improved simulation of the MJO could potentially improve long-term (climate) simulations of tropical cyclones with deep convective feedback. An example of inter-comparing high-resolution model simulations with Satellite measurements is also shown to assure the detection in TC formation.

Our vision is that the ultimate integrated modeling and CV system will enable researchers, policy and decision makers, and educators to monitor (zoom in/out) global model simulations at a wide range of spatial and temporal resolutions in real time.

Acknowledgement:

We'd like to thank the following organizations for their support: NASA Earth Science Technology Office (ESTO); Advanced Information Systems Technology (AIST) Program; NSF Science and Technology Center; NASA Modeling, Analysis Prediction (MAP) Program; the Energy and Water Cycle Study (NEWS); the NASA High-End Computing (HEC) Program, and the NASA Advanced Supercomputing (NAS) facility at Ames Research Center, and the NASA Center for Computational Science (NCCS) at Goddard Space Flight Center.

References:

Atlas, R., O. Reale, B.-W. Shen, S.-J. Lin, J.-D. Chern, W. Putman, T. Lee, K.-S. Yeh, M. Bosilovich, and J. Radakovich, 2005: Hurricane forecasting with the high-resolution NASA finite volume general circulation model. *Geophys. Res. Lett.*, **32**, L03801, doi:10.1029/2004GL021513.

Biswas, R., M.J. Aftosmis, C. Kiris, and B.-W. Shen, 2007: Petascale Computing: Impact on Future NASA Missions. Petascale Computing: Architectures and Algorithms 29-46 (D. Bader, ed.), Chapman and Hall / CRC Press, Boca Raton, FL.

Ellsworth D., B. Green, C. Henze, P. Moran, T. Sandstrom, 2006: Concurrent Visualization in a Production Supercomputing Environment. IEEE Transactions on Visualization and Computer Graphics, VOL. 12, NO. 5, September/October 2006.

Green, B., C. Henze, B.-W. Shen, 2010: Development of a scalable concurrent visualization approach for high temporal- and spatial-resolution models. AGU 2010 Western Pacific Geophysics Meeting, Taipei, Taiwan, June 22-25, 2010. (submitted)

Kerr, R., 2006: Sharpening Up Models for a Better View of the Atmosphere. *Science*, **31**, 1040.

Lander, M. 1990: Evolution of the cloud pattern during the formation of tropical cyclone twins symmetrical with respect to the equator. *Mon. Wea. Rev.*, **118**, 1194-1202.

Lang, S., W.-K. Tao, J. Simpson and B. Ferrier, 2003: Modeling of convective-stratiform precipitation processes: Sensitivity to partitioning methods, *J. Appl. Meteor*. **42**, 505-527.

Lin, S.-J., B.-W. Shen, W. P. Putman, J.-D. Chern, 2003: Application of the high-resolution finite-volume NASA/NCAR Climate Model for Medium-Range Weather Prediction Experiments. EGS - AGU - EUG Joint Assembly, Nice, France, 6 - 11 April 2003.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293-2307.

Matsui, T., X. Zeng, W.-K. Tao, H. Masunaga, W. S. Olson, and S. Lang, 2008: Evaluation of long-term cloud-resolving model simulations using satellite radiance observations and multi-frequency satellite simulators. *JAOT*.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702-708.

Maloney, E. D., and D. L. Hartmann, 2000: Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian oscillation. *Science*, **287**, 2002-2004.

NRC Decadal Survey, 2007: *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond.* The National Academies Press, Washington, D.C.

Putman, W., S.-J. Lin, and B.-W. Shen, 2005: Cross-Platform Performance of a Portable Communication Module and the NASA Finite Volume General Circulation Model International Journal of High Performance Computing Applications. 19: 213-223.

Randall D, M. Khairoutdinov, A. Arakawa, W. Grabowski, 2003: Breaking the Cloud Parameterization Deadlock. *Bull. Amer. Meteor. Soc.*, 1547-1564.

Shen, B.-W., R. Atlas, O. Reale, S.-J. Lin, J.-D. Chern, J. Chang, C. Henze, and J.-L. Li, 2006a: Hurricane Forecasts with a Global Mesoscale-Resolving Model: Preliminary Results with Hurricane Katrina (2005). *Geophys. Res. Lett.*, 33, L13813, doi:10.1029/2006GL026143.

Shen, B.-W., R. Atlas, J.-D. Chern, O. Reale, S.-J. Lin, T. Lee, and J. Chang, 2006b: The 0.125 degree finite-volume General Circulation Model on the NASA Columbia Supercomputer: Preliminary Simulations of Mesoscale Vortices. *Geophys. Res. Lett.*, **33**, L05801, doi:10.1029/2005GL024594.

Shen, B.-W., W.-K. Tao, R. Atlas, T. Lee, O. Reale, J.-D. Chern, S.-J. Lin, J. Chang, C. Henze, J.-L. Li, 2006c: Hurricane Forecasts with a Global Mesoscale-resolving Model on the NASA Columbia Supercomputer, AGU 2006 Fall Meeting, December 11-16, 2006.

Shen, B.-W. and W.-K. Tao, and R. Atlas, 2009a: Hurricane Forecasts with the Global Mesoscale model on NASA Supercomputers. International Conference for High-Performance Computing in ASIA-Pacific Region 2009. Proceedings, p649-650. Kaohsiung, Taiwan, March 2-5, 2009. This is available from http://atmospheres.gsfc.nasa.gov/cloud_modeling/docs/Shen_poster_HPC_ASIA_2009_Proceedings.pdf

Shen, B.-W., W.-K. Tao, G. Bryan, C. Henze, P. Mehrotra, J.-L. F. Li, S. Cheung, 2009b: High-impact Tropical Weather Prediction with the NASA CAMVis: Coupled Advanced multi-scale Modeling and concurrent Visualization Systems. Supercomputing conference 2009, Portland, Oregon, November 14-20, 2009. This is available at http://atmospheres.gsfc.nasa.gov/cloud_modeling/docs/SC09_Shen_Poster.pdf and ftp://meso.gsfc.nasa.gov/betsy/shen/resume/docs/SC09_Shen_Movie.avi (movie; ~60MB)

Shen, B.-W., W.-K. Tao, J.-D. Chern, and R. Atlas, 2009c: Scalability Improvements in the NASA Goddard Multiscale Multicomponent Modeling Framework for Tropical Cyclone Climate Studies. International Conference for High-Performance Computing in ASIA-Pacific Region 2009. Proceedings, p249-256. Kaohsiung, Taiwan, March 2-5, 2009. This is available from http://atmospheres.gsfc.nasa.gov/cloud_modeling/docs/Shen_HPC_ASIA_2009_Proceedings.pdf

Shen, B.-W., W.-K., Tao, W. K. Lau, and R. Atlas, 2010a: Improving Tropical Cyclogenesis Prediction with a Global Mesoscale Model: Hierarchical Multiscale Interactions During the Formation of Tropical Cyclone Nargis (2008). J. Geophys. Res., doi:10.1029/2009JD013140.

Shen, B.-W., W.-K. Tao, R. Atlas, Y.-L. Lin, C.-D. Peters-Lidard, J.-D. Chern, K.-S. Kuo, 2010b: Forecasting Tropical Cyclogenesis with a Global Mesoscale Model: Preliminary Results for Twin Tropical Cyclones in May 2002. (to be submitted)

Taft, J. R., 2001: Achieving 60 gflop/s on the production cfd code overflow-mlp. *Parallel Computing*, 27(4):521-536.

Tao, W.-K., and J. Simpson, 1993: The Goddard Cumulus Ensemble Model. Part I: Model description. *Terrestrial, Atmospheric and Oceanic Sciences*, **4**, 19-54.

Tao, W.-K., and J. Simpson, C.-H. Sui, B. Ferrier, S. Lang, J. Scala, M.-D. Chou, and K. Pickering, 1993: Heating, moisture and water budgets of tropical and mid-latitude squall lines: Comparisons and sensitivity to longwave radiation. *J. Atmos. Sci.*, **50**, 673-690.

Tao, W.-K., D. Anderson, R. Atlas, J. Chern, P. Houser, A. Hou, S. Lang, W. Lau, C. Peters-Lidard, R. Kakar, S. Kumar, W. Lapenta, X. Li, T. Matsui, R. Rienecker, B.-W. Shen, J. J. Shi, J. Simpson, and X. Zeng, 2008: A Goddard Multi-Scale Modeling System with Unified Physics. WCRP/GEWEX Newsletter, Vol 18, No 1, 6-8.

Tao, W.-K., J. Chern, R. Atlas, D. Randall, X. Lin, M. Khairoutdinov, J._L. Li, D. E. Waliser, A. Hou, C. Peters-Lidard, W. Lau, and J. Simpson, 2009b: Multi-scale modeling system: Development, applications and critical issues, *Bull. Amer. Meteor. Soc.* **(**March Issue**).**

Tao, W.-K., J. Shi, S. Chen, S. Lang, S.-Y. Hong, G. Thompson, C. Peters-Lidard, T. Matsui, A. Hou, S. Braun, and J. Simpson, 2009a: Studying precipitation processes in WRF with Goddard bulk microphysics: Comparison with different microphysical schemes,, *Mon. Wea. Rev.*, **(**submitted**).**