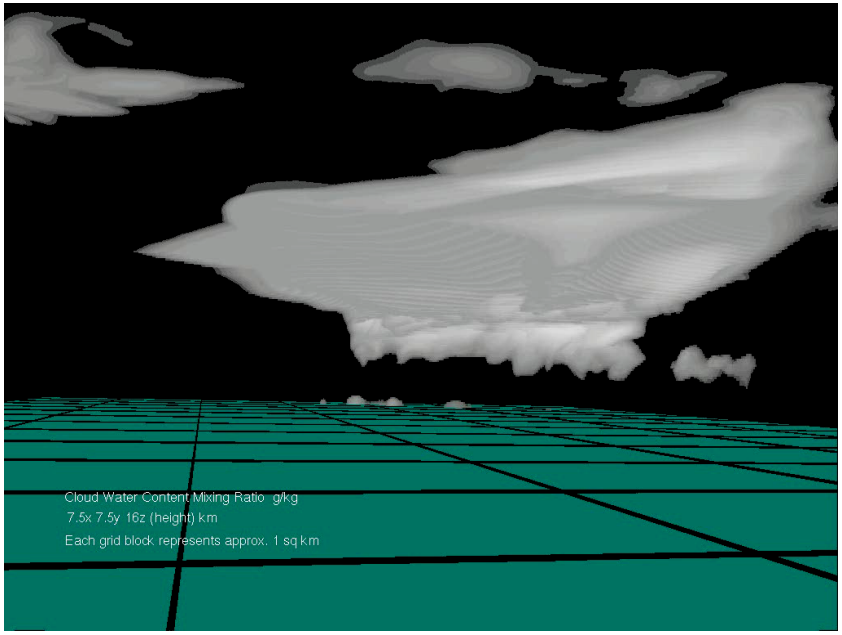


Tornado (25 m grid spacing)

Prof. Ming Xue at U. Oklahoma

Hurricane Katrina (2005)

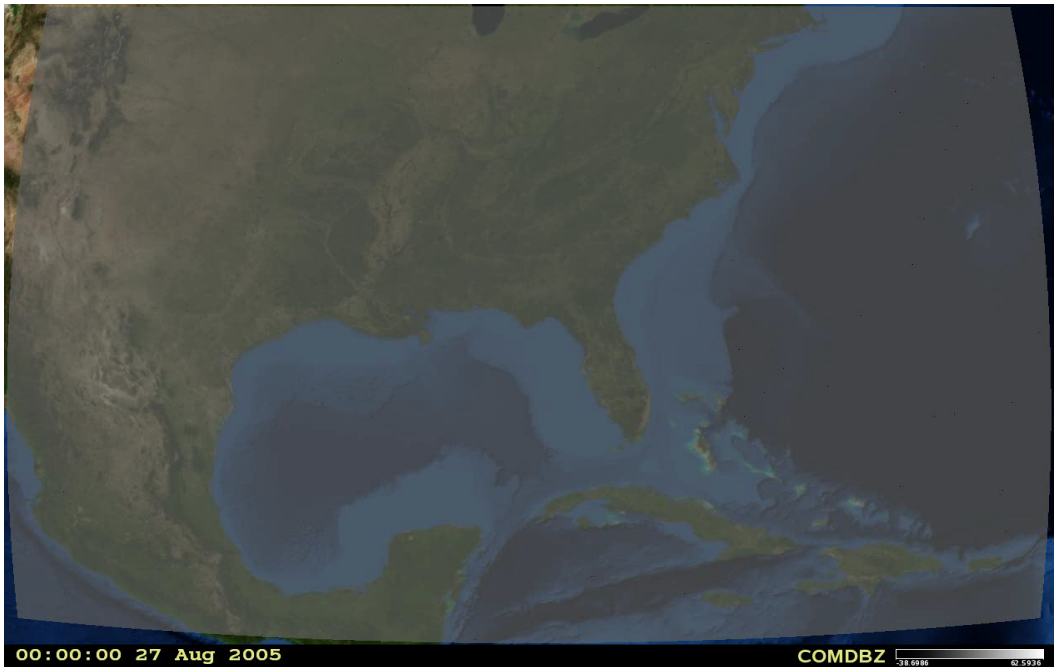
1.67 km grid spacing

18 second loop

Tao et al. (2011)

Ames

Microphysics
Resolution
Initial Condition



Empowering Data Management, Diagnosis, and Visualization of *Cloud-Resolving Models* (CRM) by Cloud Library upon Spark and Hadoop

Wei-Kuo TAO (PI, GSFC)
Xian-He Sun (Co-PI, IIT)
Shujia Zhou (Co-PI, Northrop Grumman)
Toshihisa Matsui (Co-PI, ESSIC UMD)
Xiaowen Li (Co-I, GESTAR MSU)
Dan Duffy (Collaborator, GSFC)

Outlines →

Background/Goals/Approach
Modeling Systems
Pilot Cases
Results
Status

ESTO Earth Science Technology Office

Goddard Cloud Library

Mesoscale Dynamic and Modeling:
http://cloud.gsfc.nasa.gov/

Model descriptions, Group members, Publications & multi-dimensional (space, time, multivariate, and multiple cloud/cloud system type) cloud database (model simulated cloud data) is available from a web site created within NASA Goddard (Goddard Cloud Library).

Currently, more than 70,000 datasets were downloaded since April 2010 by155 distinct users

Modeled + Observed Data in Mass Storage
No visualization nor diagnosis

Supported by NASA MAP

ESTO
Earth Science Technology Office

Goals

- Make Cloud Resolving Model output more usable by science community
  - Accelerate visualization of output.
  - Inter-compare large volumes of output from high-resolution simulations.
  - Diagnose key processes for cloud-precipitation.

- Demonstrate the value to distribute, visualize, analyze and inter-compare Cloud Resolving Model output and data with GCE and NU-WRF

GCE: Goddard Cumulus Ensemble

NU-WRF: NASA Unified Weather Research Forecast

Approach

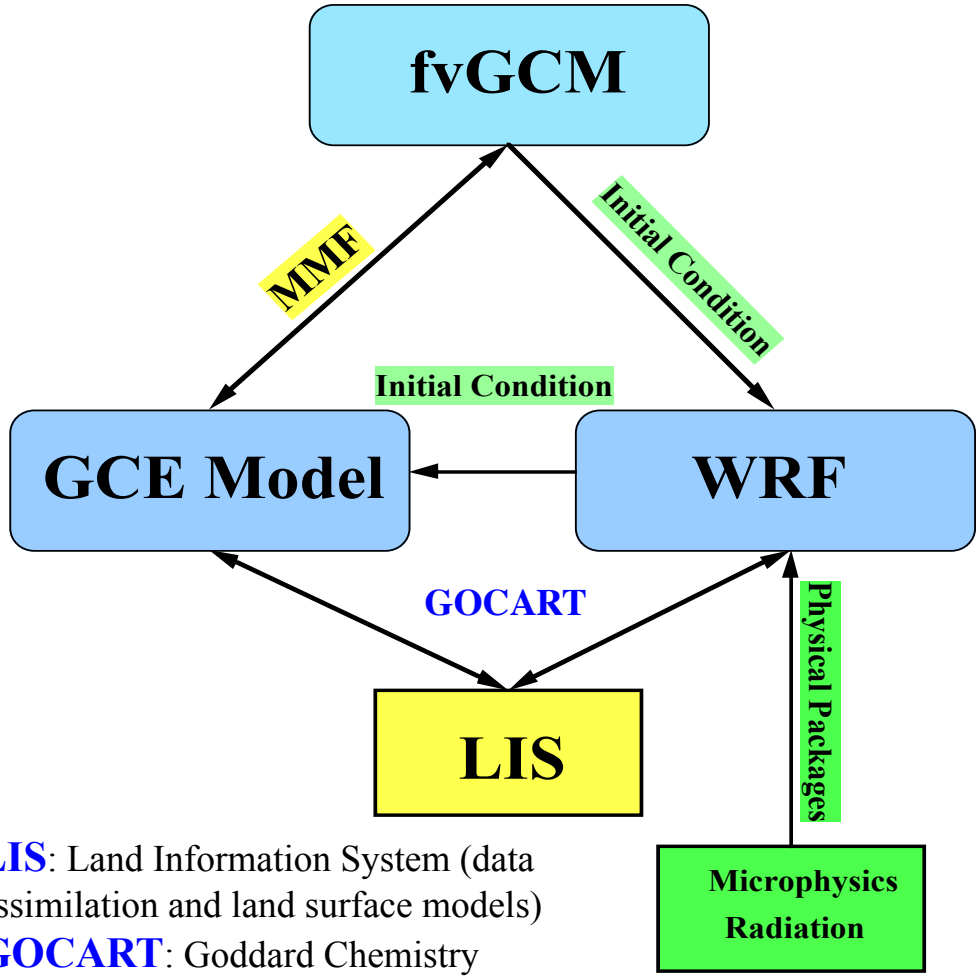
- Develop Super Cloud Library (SCL) supporting Cloud Resolving Model using Spark on Hadoop.
  - *Create cloud data files*
  - *Develop data model and Hadoop format transformer*
  - *Develop dynamic transfer tool to Hadoop*
  - Develop subset and visualization APIs (Application Programming Interfaces)
  - Develop Web User Interface
- Conduct **Demo** of GCE and NU-WRF diagnosis on NAS and NCCS

February 15, 2015

Blue: Have preliminary results

NASA Cloud Resolving Models

- **Multi-scale modeling system developed at Goddard with unified physics** from:
  1. **Goddard Cumulus Ensemble model** (GCE), a cloud-resolving model (CRM)
  2. **NASA unified Weather Research and Forecasting Model** (WRF), a region-scale model, and
  3. Coupled fvGCM-GCE, the GCE coupled to a general circulation model (or GCM known as **Goddard Multi-scale Modeling Framework or MMF**).
- Same parameterization schemes all of the models for cloud microphysical processes, long- and short-wave radiative transfer, and land-surface processes, to study explicit cloud-radiation and cloud-surface interactive processes.
- ***Coupled with multi-sensor simulators for comparison and validation of NASA high-resolution satellite data.***



LIS: Land Information System (data assimilation and land surface models)
GOCART: Goddard Chemistry Aerosol Radiation and Transport Model

Tao, W.-K., S. Lang, X. Zeng, X. Li, T. Matsui, K. Mohr, D. Posselt, J. Chern, C. Peters-Lidard, P. Norris, I.-S. Kang, A. Hou, K.-M. Lau, I. Choi, M. Yang, 2014: The Goddard Cumulus Ensemble (GCE) Model: Improvements and Applications for Studying Precipitation Processes. An invited paper - *Atmos. Res.*, **143**, 392-424

ESTO Earth Science Technology Office

Goddard Cumulus Ensemble (GCE) Model
(1982 – , over 150 papers published in last 3 decades)

Water and energy cycles in the tropical climate system,

Redistribution of ozone and trace,

Deep convection related to global warming,

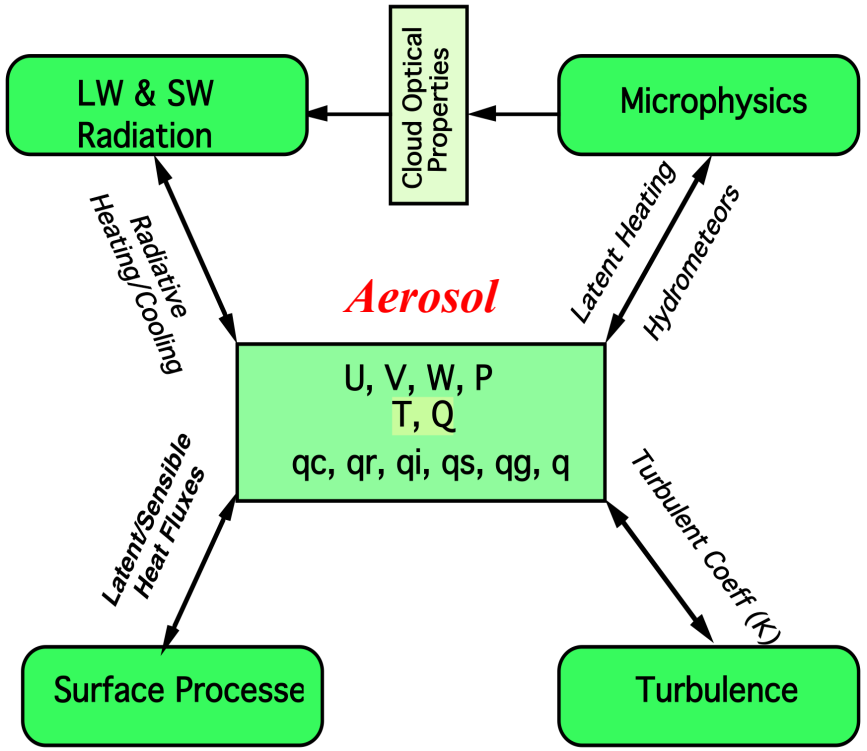
Precipitation processes - precipitation efficiency,

The aerosol impact on precipitation and rainfall,

Surface process on precipitation and rainfall,

Representation of cloud microphysical processes and their interaction with radiative forcing

TRMM and GPM rainfall and latent heating retrieve algorithm performance



GCE Model Description: Tao and Simpson (1993), Tao et al. (2003), Tao (2003), **Tao et al. (2014)**
CRM review paper: Tao and Moncrieff (1999 – Geophy Review)
Aerosol review paper: Tao et al. (2012 – Geophy Rev)

Supported by NASA TRMM/GPM and MAP

ESTO Earth Science Technology Office

NU-Unified WRF

Peters-Lidard *et al.* (2014)

Tao *et al.* (2013, 2015), Matsui *et al.* (2014)

Real time forecast to support GPM ground validation field campaigns

Precipitation processes study to support satellite missions (GPM and CaPPM)

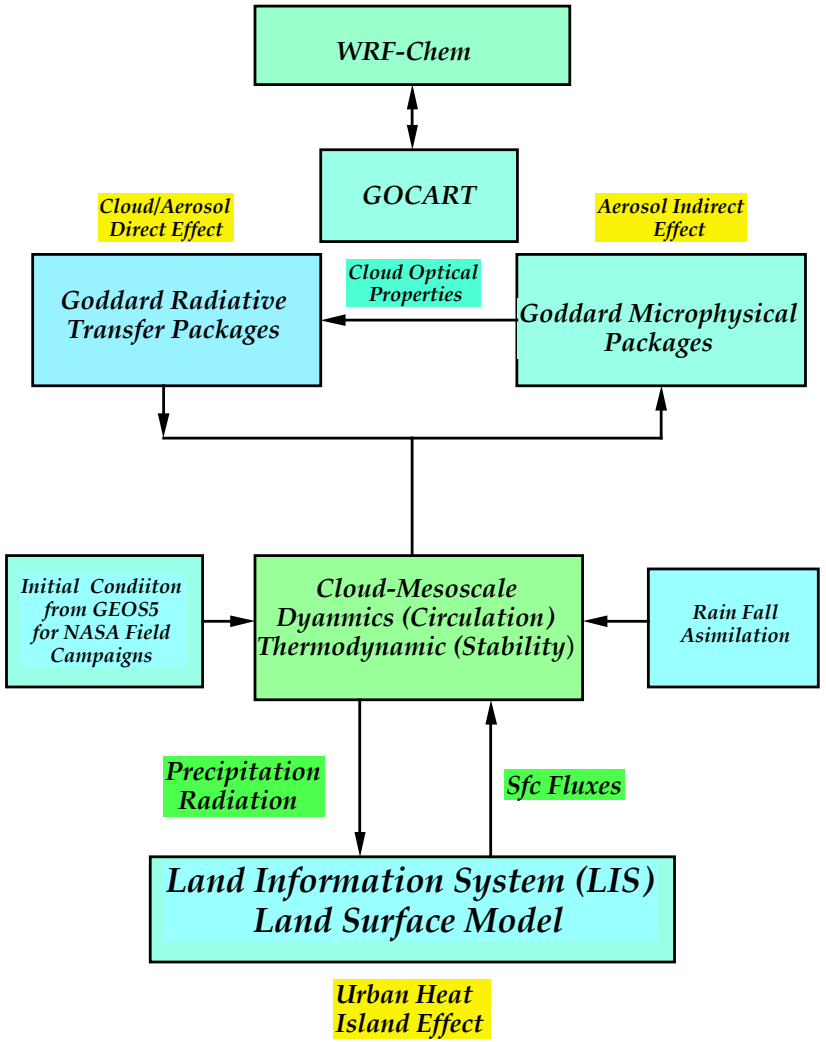
Cloud-aerosol-precipitation interaction (direct and indirect)

Regional climate (downscaling from climate models)

Dust effect

Land surface effect on precipitation

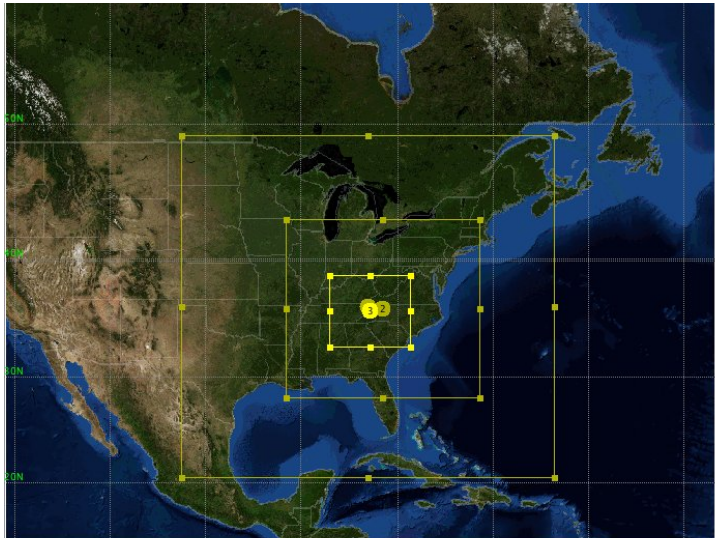
Hydrological study (flood, draught)



Blue box:
Goddard Physical Packages

Supported by NASA MAP

ESTO Earth Science Technology Office



NU-WRF : IPHEX

(May 1 – June 15, 2014)

9km(386x353), 3km(601x553), **1km(751x667)**

Computational Cost: **2048 CPUs, takes 7 hours to produce 48 hours forecast.**

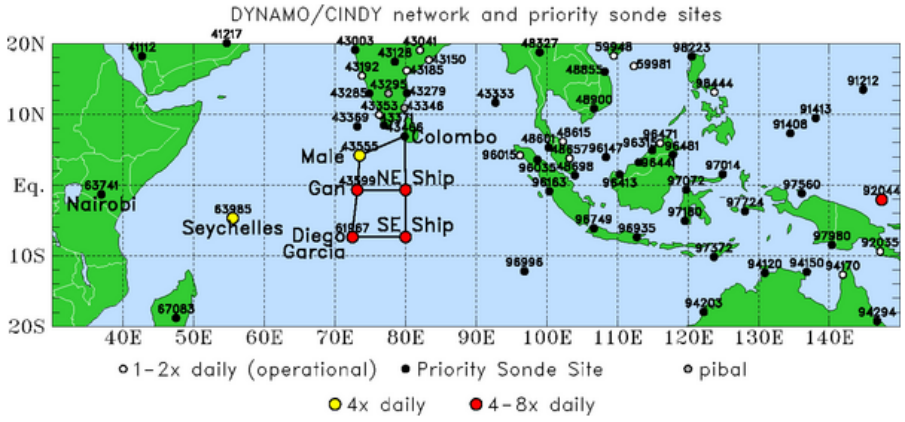
Long-Term: AMMA case (Phase I &III):

Phase I Jan & Feb (Dry season: aerosol dominate)

Phase III July-Aug (Peak Monsoon: MCS).

Short-Term: 1-0.5 km Nature run:

April 27–30, 2014 tornado outbreak

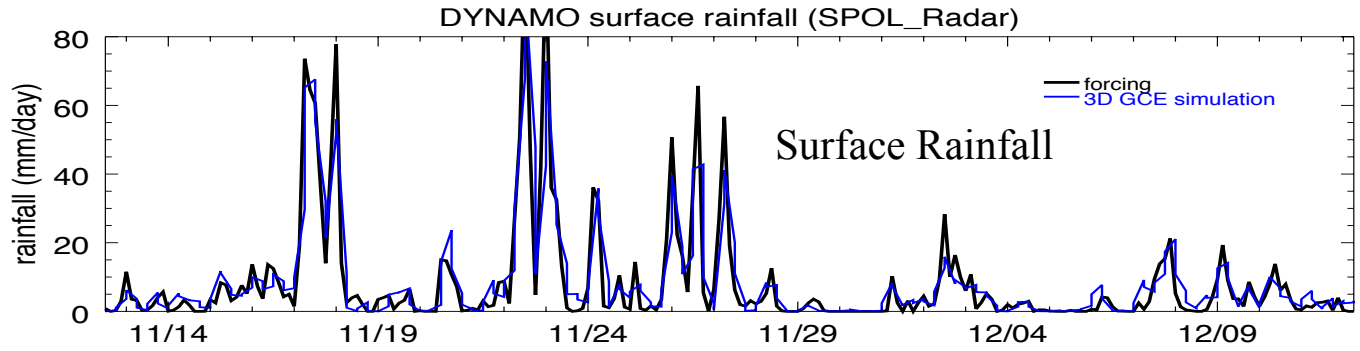


GIGA GCE-LES

4096x4096x104, **250 m**, 2-3 day simulation:

Total output data size ~125TB

Model Inter-comparison

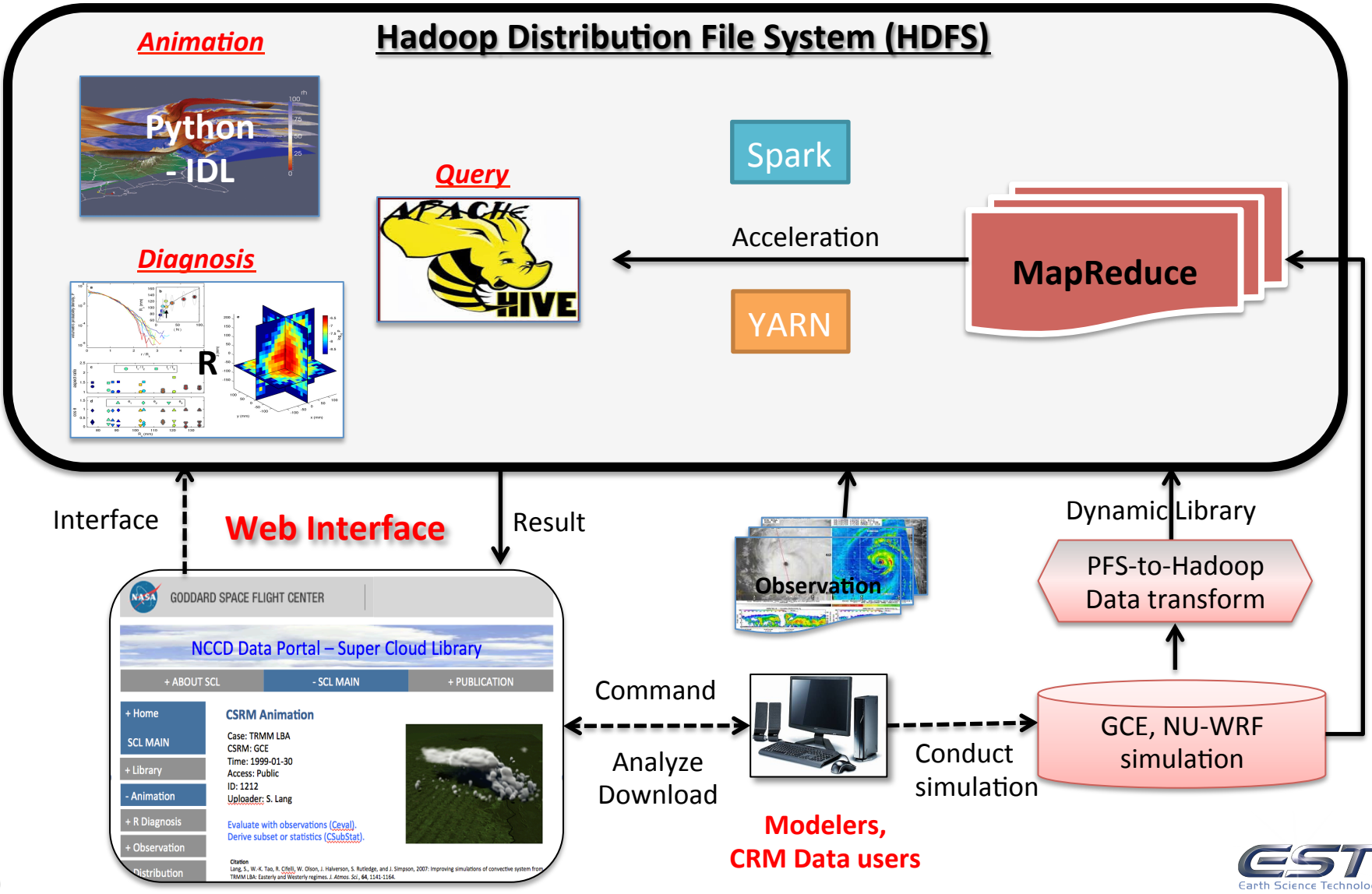


DYNAMO MJO

Nov 10 – Dec 14

ESTO Earth Science Technology Office

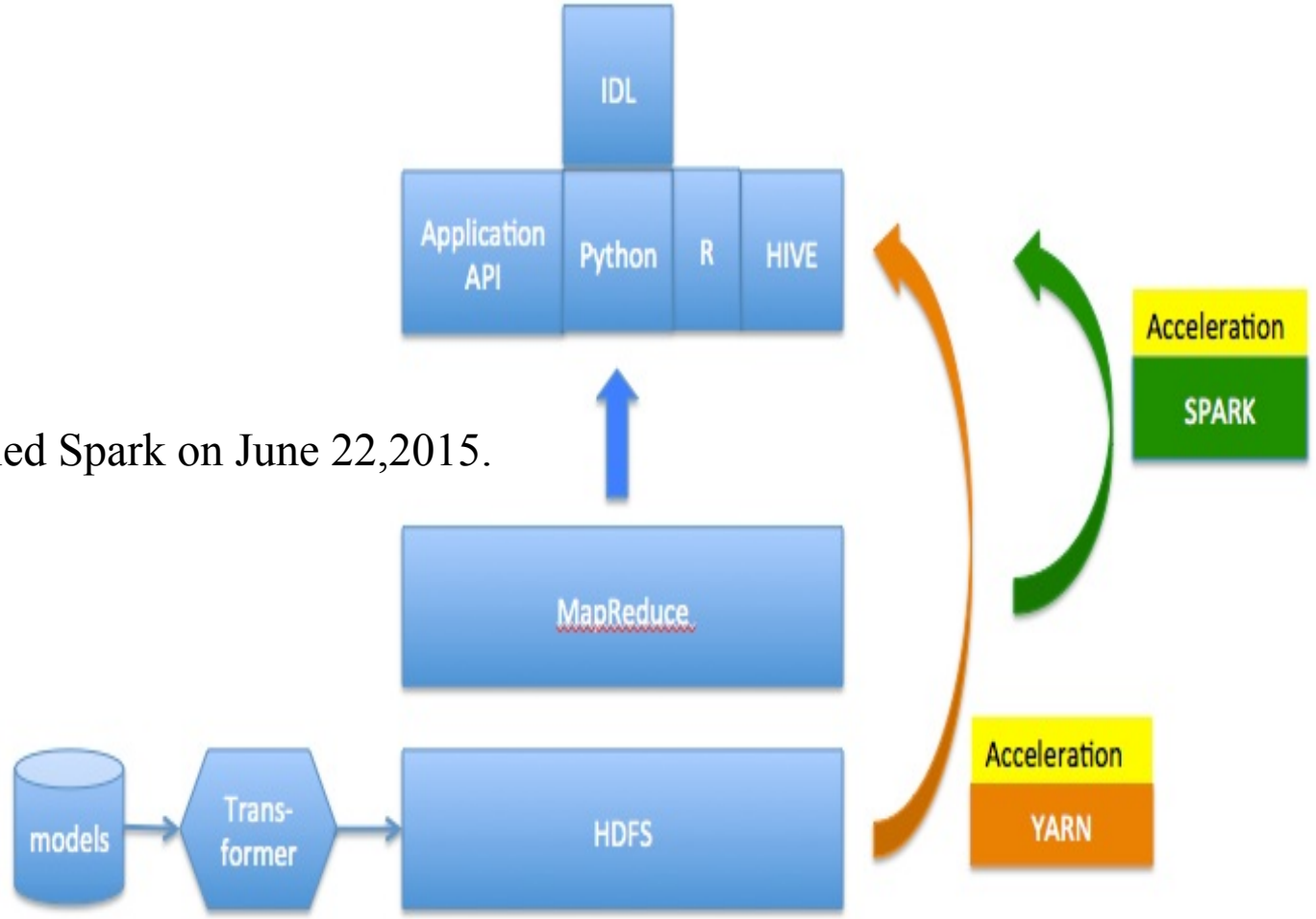
Super Cloud Library (SCL)



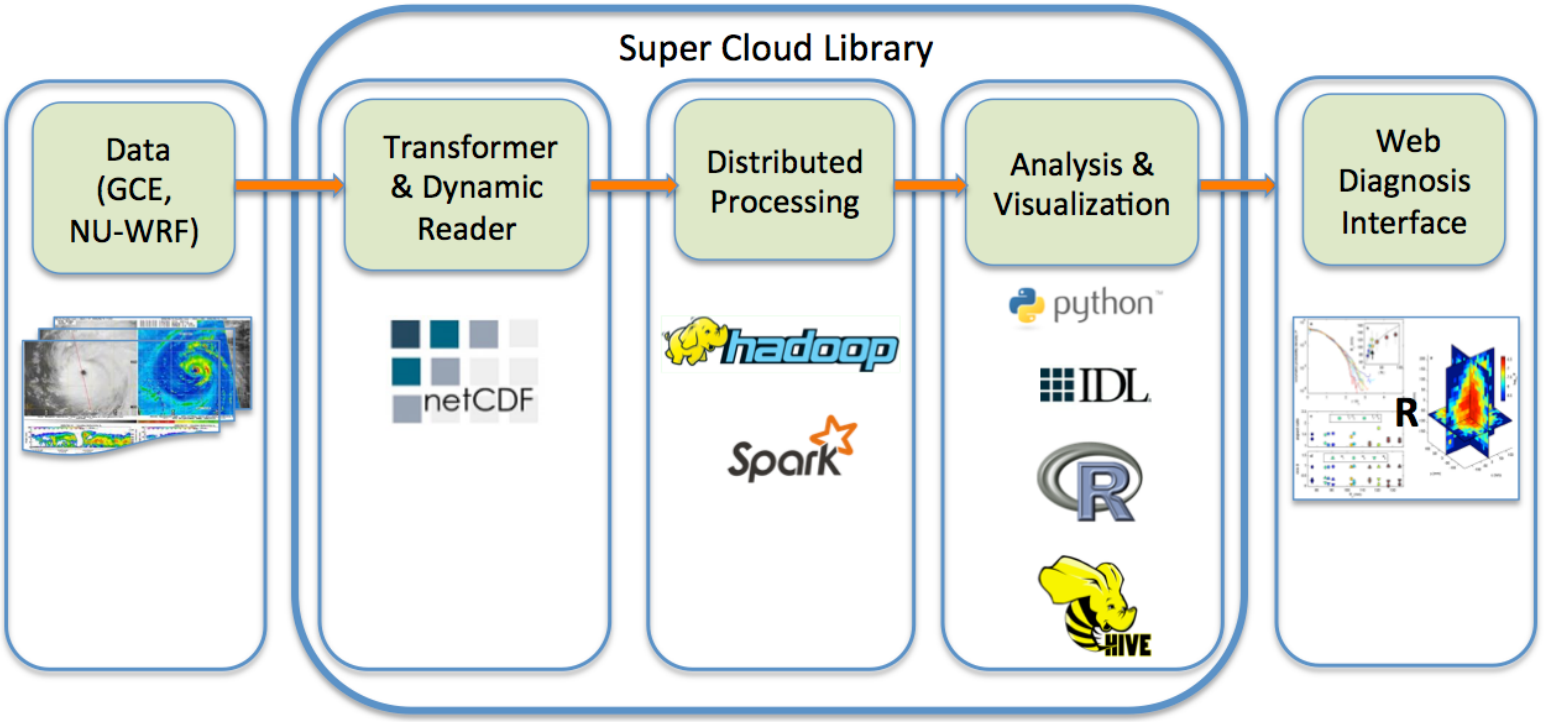
ESTO
Earth Science Technology Office

Illustration of super cloud library from the viewpoint of applications

NCCS just installed Spark on June 22,2015.



Hadoop Distributed File System (HDFS), Yet Another Resource Negotiator (YARN)
Spark uses memory for acceleration and YARN use resource management for acceleration
HIVE is a SQL-like language/tool for Hadoop. Impala is an optimized version of HIVE
Python/R is a script language/tool supported by Hadoop and Spark.



The difference between ***pilot and proposed*** case is the size of model simulation and data storage.

The pilot case allows us to develop and test data model as well as all software (e.g., Hadoop, PFS, IDL, R)

PFS: Parallel File System
R: A popular programming language and software environment for statistical computing and graphics

Data Sizes (pilot vs proposed cases)

- Pilot case

  GCE (250x250x40, month-long simulation): *DYNAMO Nov. 2011 MJO event*; Single output data size 3~6GB; Total output data size ~2TB.

  NU-WRF (1 km and 48 h simulation): IPHEX; 10 output variables. Hourly output in a text (CSV) format has 465GB

- 1KMSBM, 250x250x40, 5-day simulation: Single output data size ~60GB; Total output data size ~7.2TB.

- GIGALES, 4096x4096x104, 2-3 day simulation: Single output data size ~2.5TB; Total output data size ~125TB.

Data Model

- Read NU-WRF NetCDF output or GCE binary output, write it out in a text (CSV) format and provide a file with the relevant meta data (e.g., unit and time stamp) for reconstructing it into CF-NetCDF
  - CF-NetCDF is a NetCDF with CF compliance
- One variable, one time frame per file
  - One variable, multiple time frames per file
- A subset data file will be available in CF-NetCDF

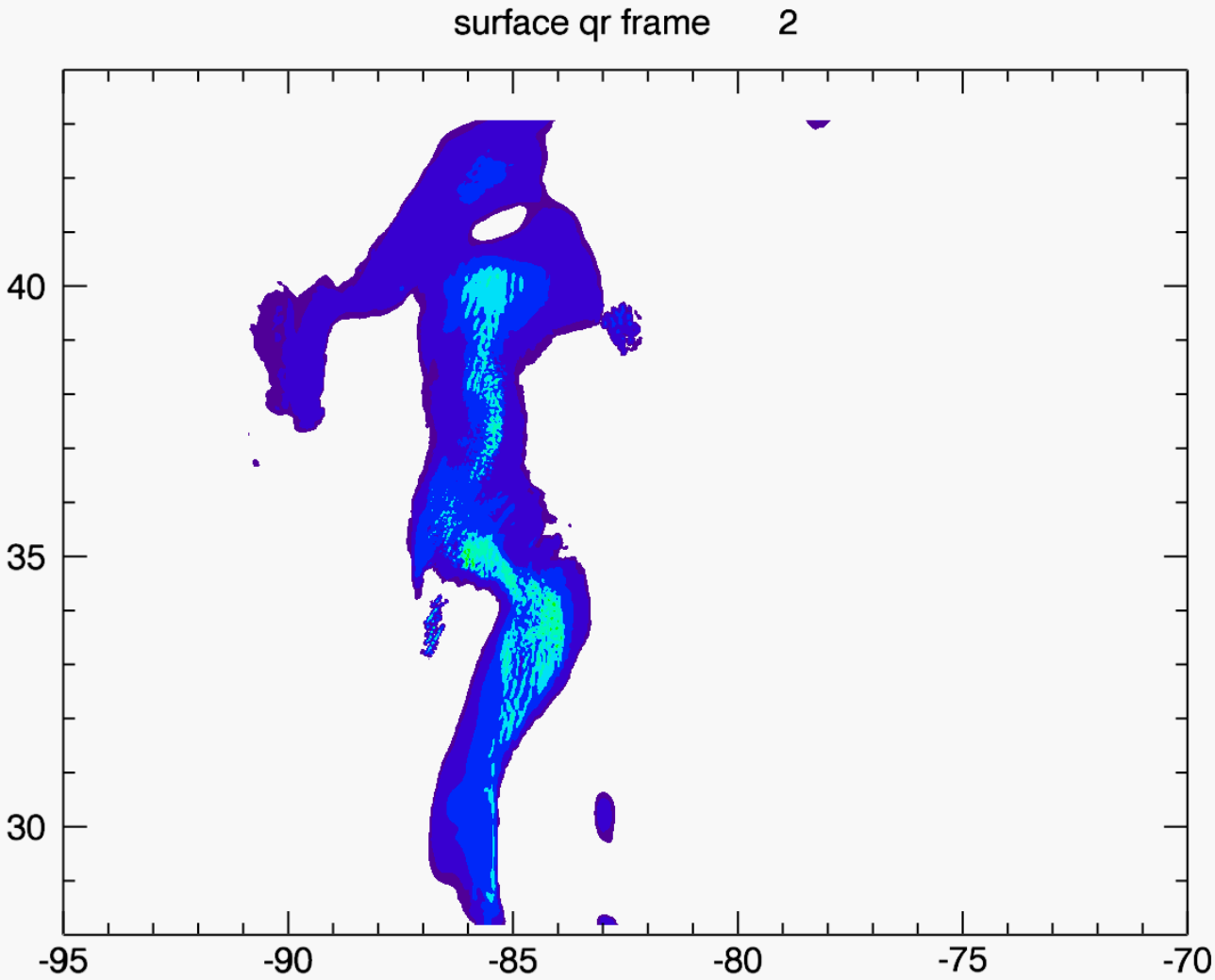
CSV: Comma-separated value
CF: Climate and Forecast Metadata

IDL+Hadoop Performance

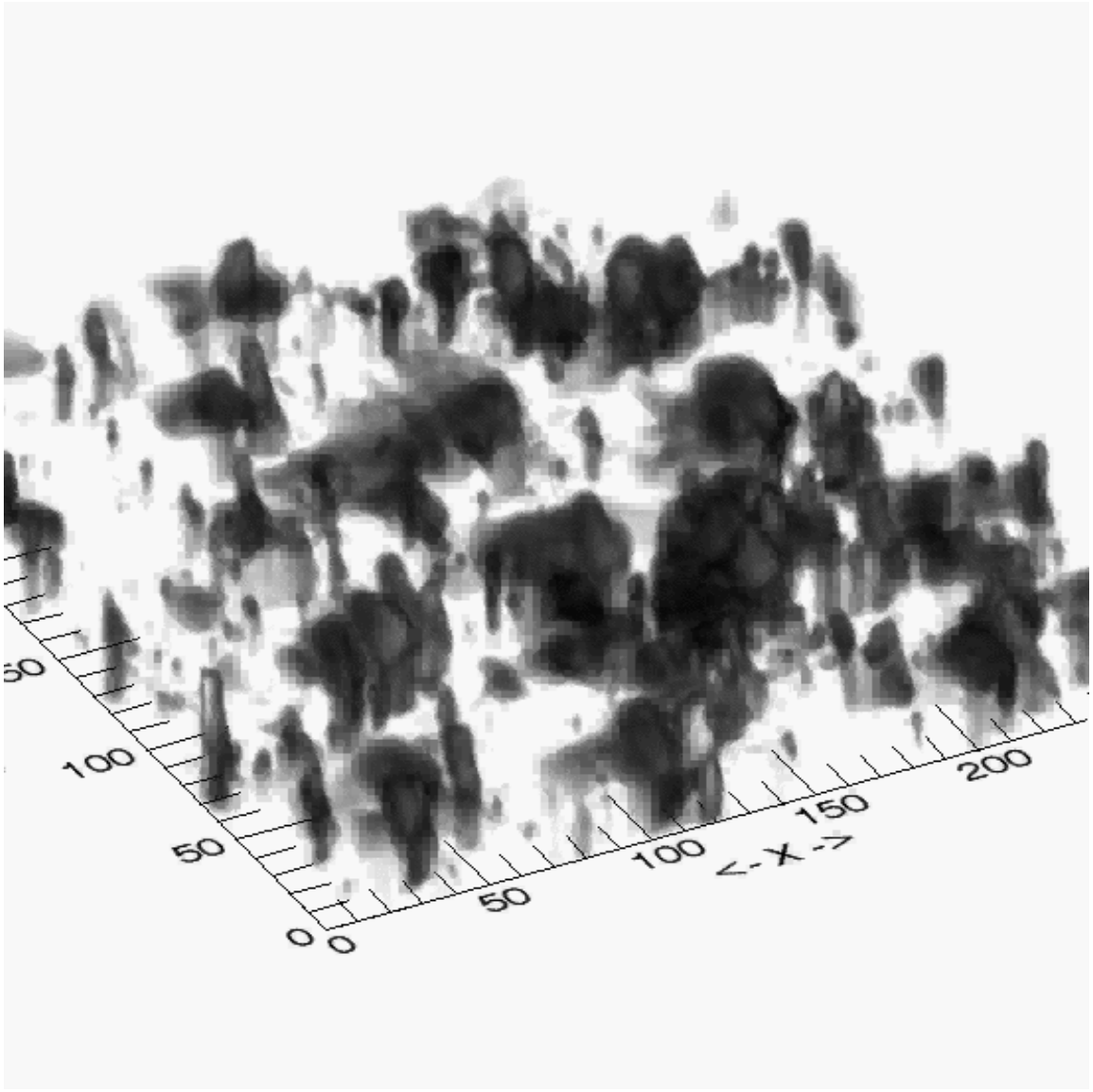
- Developed an animation/movie of 48 text files from NU-WRF Pilot simulations. Each file has ~1GB
  - 14 minutes for IDL with Hadoop streaming
    - NCCS Hadoop cluster has 34 nodes
    - IDL is on one Hadoop node and can run with multiple instances simultaneously
    - Files stored in Hadoop file system (HDFS)
    - 15 seconds for reading each file out of HDFS
  - 17 minutes for IDL with NCCS Discover GPFS file system
    - IDL is on one NCCS Dali compute node
    - Files stored in NCCS GPFS file system

GIF is one image format and animation is movie
GPFS is General Parallel File System

NU-WRF
Surface Rain Water



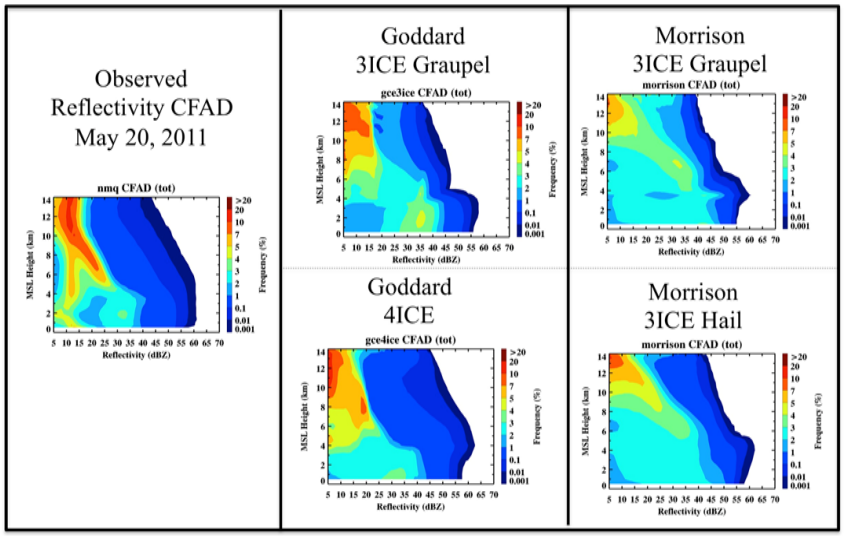
GCE 3D Cloud Structure



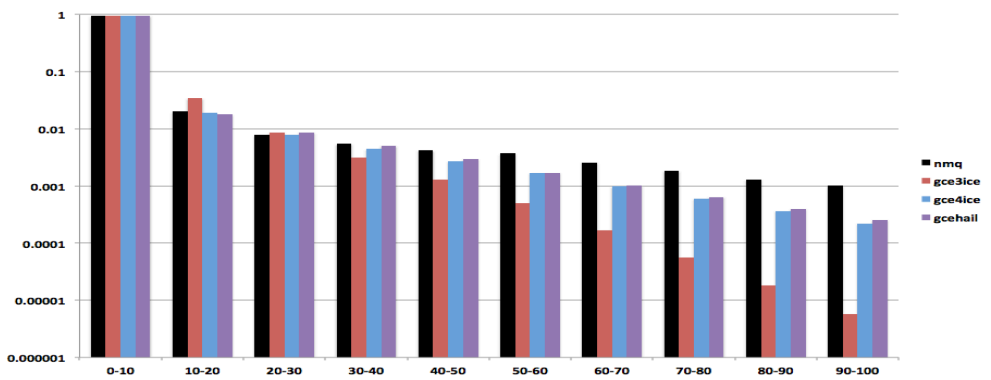
ESTO
Earth Science Technology Office

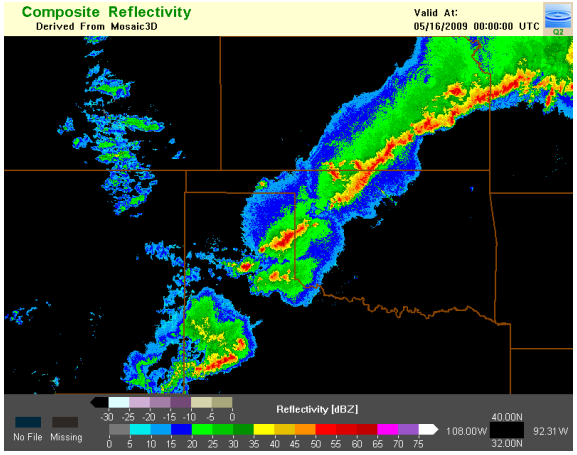
Examples of Diagnosis
(model inter-comparison)

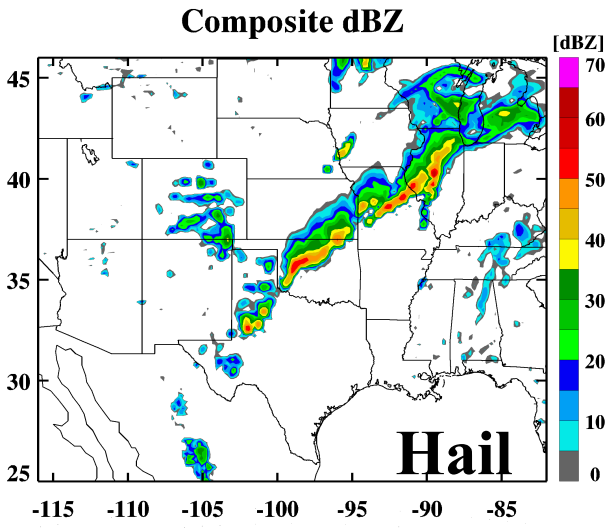
Radar Reflectivity

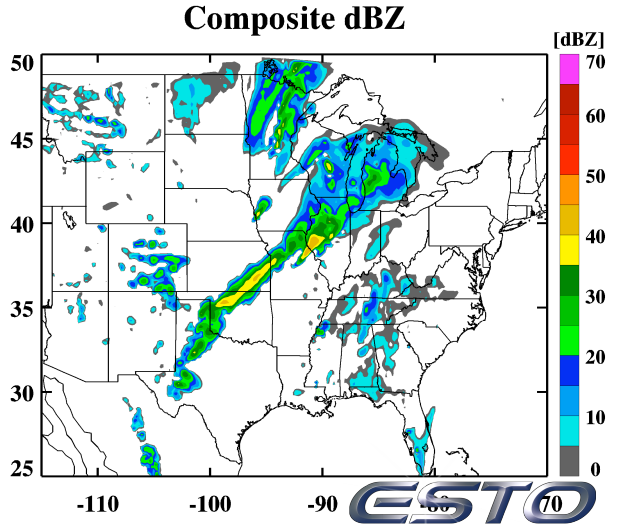


PDF – Rainfall Intensity





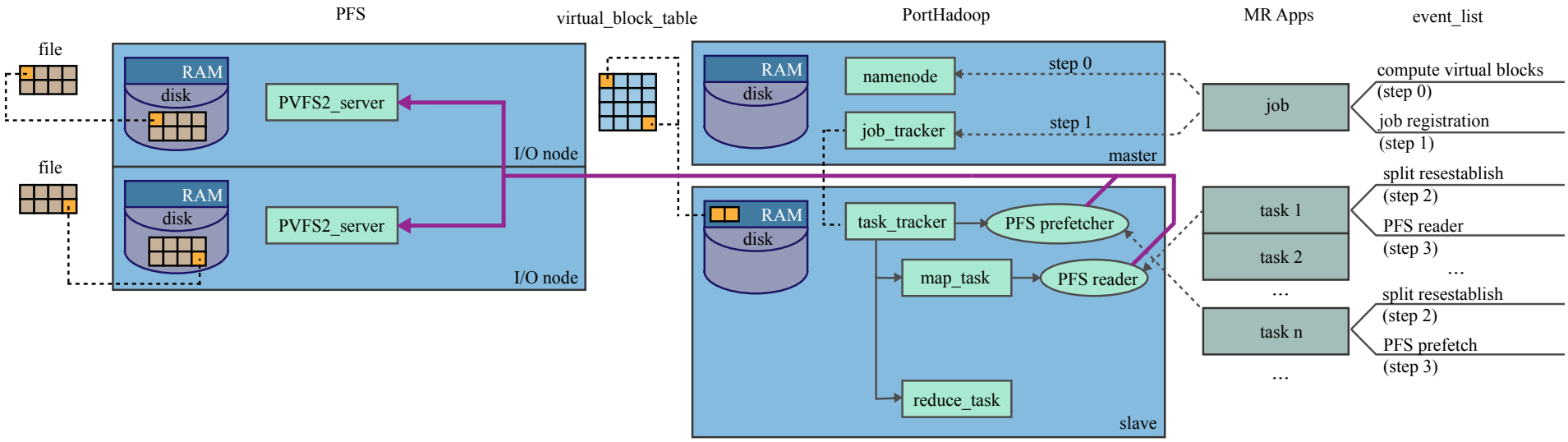




ESTO
Earth Science Technology Office

Dynamic Hadoop Reader System

- Hadoop with PVFS2 as well as CephFS parallel file systems



Job trackers and task trackers run on the master node, and task trackers and map tasks run on slave nodes. Virtual blocks are managed in the virtual block table in namenode. A task tracker initiates the PFS prefetcher thread to get data directly from a remote PFS. Map tasks initiate the PFS reader threads to import data directly from the remote PFS.

ESTO
Earth Science Technology Office

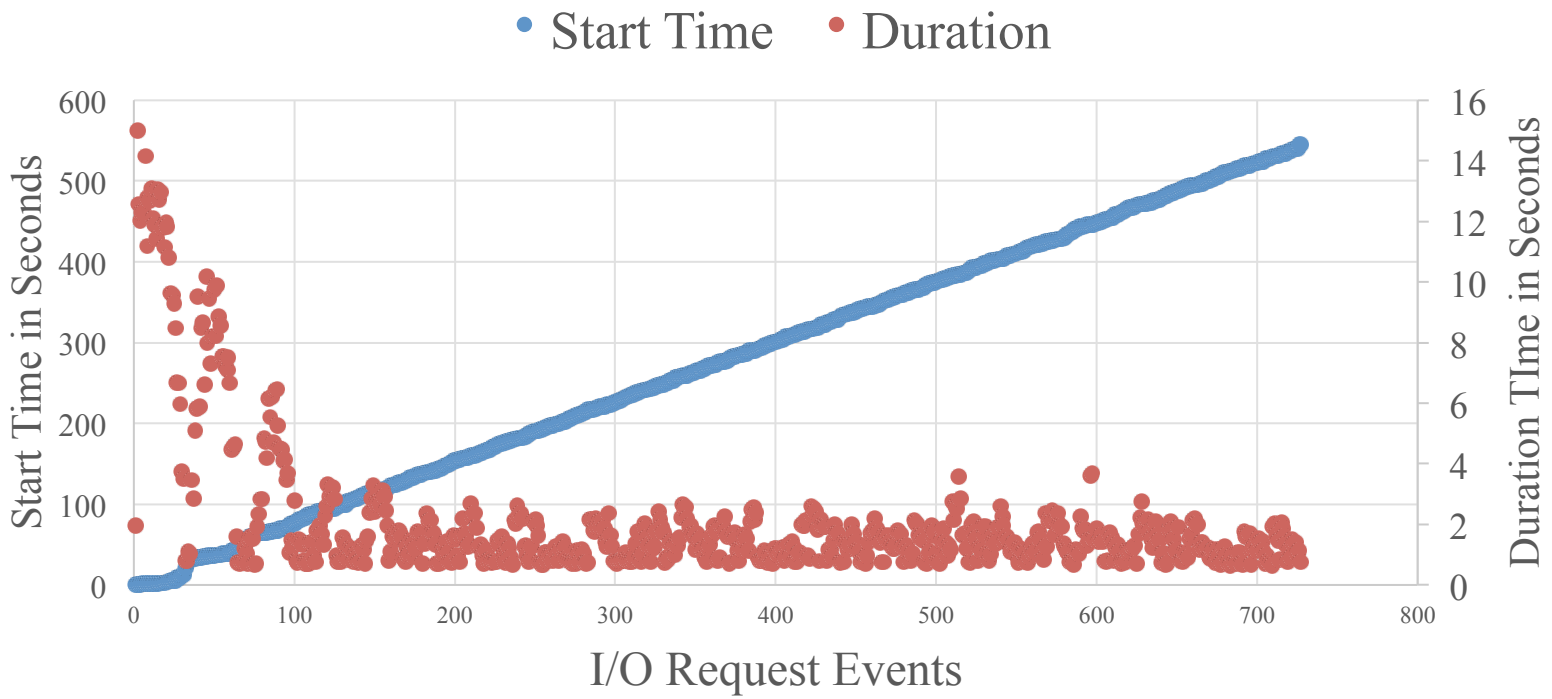
Testing Dynamic Hadoop Reader

- Test system
  - Parallel File System (PFS) can be CephFS and PVFS2
  - 1Gbit Ethernet connection
  - CephFS PFS and dynamic Hadoop reader are deployed on a 17 node cluster
  - PVFS2 FPS and dynamic Hadoop reader are deployed on a 48 node cluster
- Test case
  - Read and process with Word Count application
  - 48 text files, ~1GB each, at PVFS2
  - 8 PVFS2 I/O nodes, 8 Hadoop slave nodes
  - The block size at Hadoop is 64MB
  - 4 concurrent map tasks per node
  - Max parallelism is 32 (=4x8)

CephFS and PVFS2 are two different parallel file systems.
CephFS version 0.80.7 and OrangeFS version 2.8.6 (commercial version of PVFS2) are used

ESTO Earth Science Technology Office

Dynamic Hadoop Reader Performance



- Total time (I/O + Computation) is 561 seconds (= 9.35 minutes)
- Average I/O duration per request (64 MB) is 2.08 seconds
- Typical computation time per task is 20~25 seconds
- Longer duration time in the earlier reading requests is due to burst I/O requests
- Overlapping computation with I/O improves efficiency
- **The prototype of Dynamic Hadoop Reader is functional and being improved**

ESTO
Earth Science Technology Office

Technology

Hi: High performance, **Med:** Medium performance: **Low:** Low performance
Expected: It is expected to be from high to med.

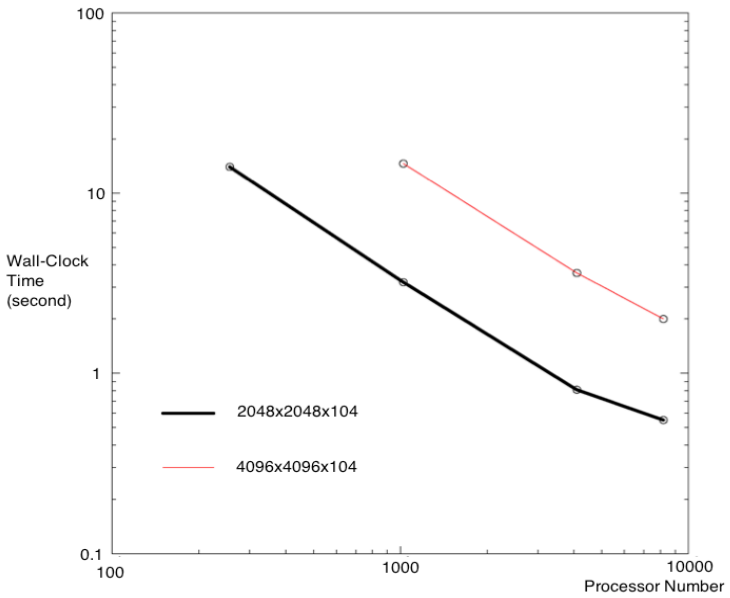
	IDL	R	HIVE	Revolution R	SparkR
Easy to use	Hi	Med	Hi	Med	Expected
Big data visualization	Hi	No	No	Hi	Expected
Big data statistical calculation	Hi	No	Med	Med	Expected
Advanced big image analysis	Low	No	No	No	Expected
Cost	$400 per node	Free	Free	$6000 per node	Free

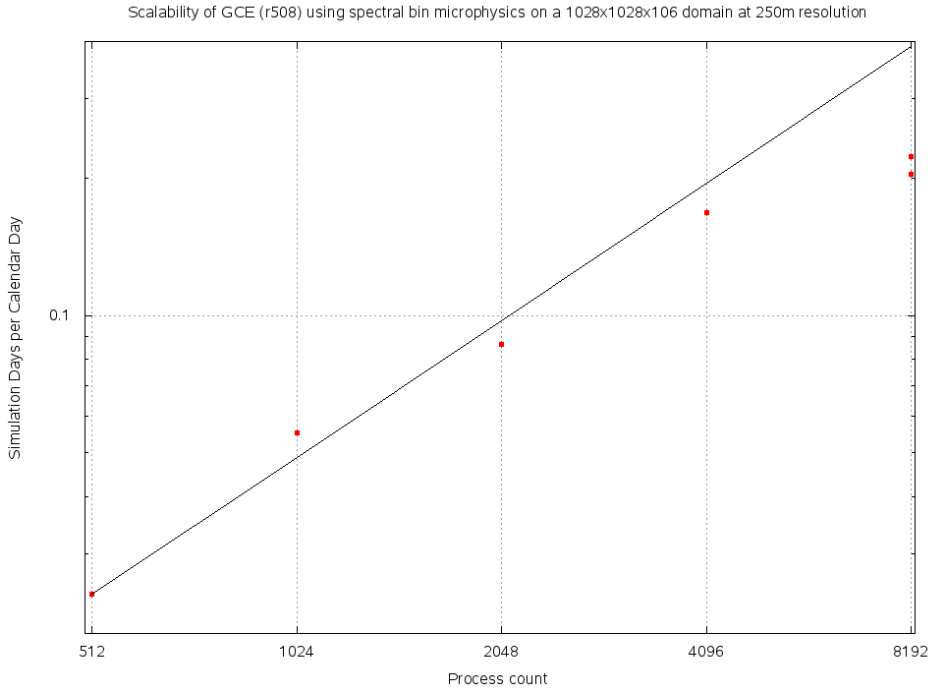
Revolution R: an enhanced version of open source R
SparkR: an R package that provides a light-weight front end to use Apache Spark from R.

Status

- Developing a data model so that various CRM simulation outputs in NetCDF and binary, including NU-WRF and GCE models, can be accessed and processed by Hadoop
- Developing a NetCDF-to-CSV converter to support NU-WRF and GCE model outputs
- Developing a technique to visualize data at Hadoop with IDL and Hadoop streaming
- Developing a use case where CRM modeled data can be processed (e.g., subsetting) with HIVE queries interactively via HUE's Web interface
  - Experimenting with Clouder Impala to speed up subsetting
- Prototyping a portable dynamic Hadoop reader to access data in a parallel file system, PVFS2 or CephFS

GCE Scalability with Bulk /without IO (left)
Spectral Bin Physics and Parallel IO (right)





As the number of processes increases, the domain size decreases and the percentage of communication cost (halo update) increase. Consequently, scalability is not linear.

However, when the domain size is fixed, it is close to liner. For example, the run with **2048x2048x104 with 4096** processes takes the similar walk clock time **to 4096x4096x104 with 1024 processes.**

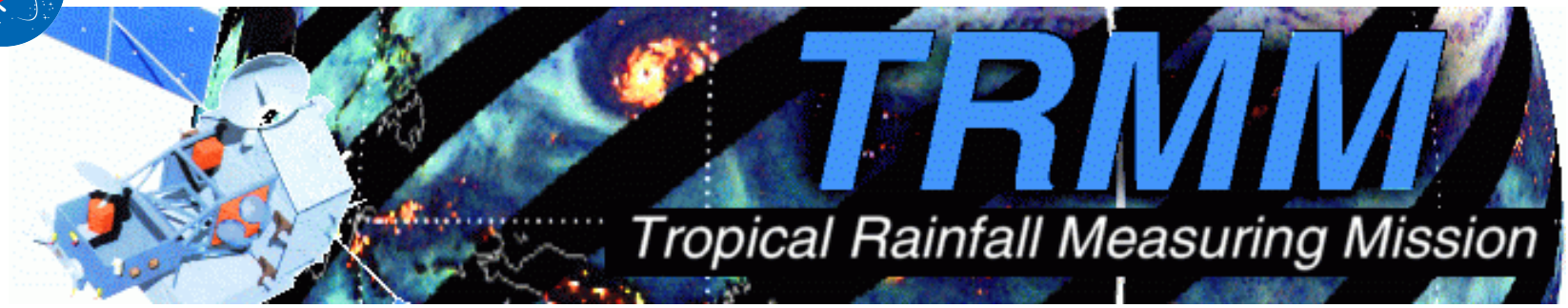
Without parallel IO, 1028x1028x46 with spectral bin microphysics run fails due to memory limitation of a single processor core.

+1

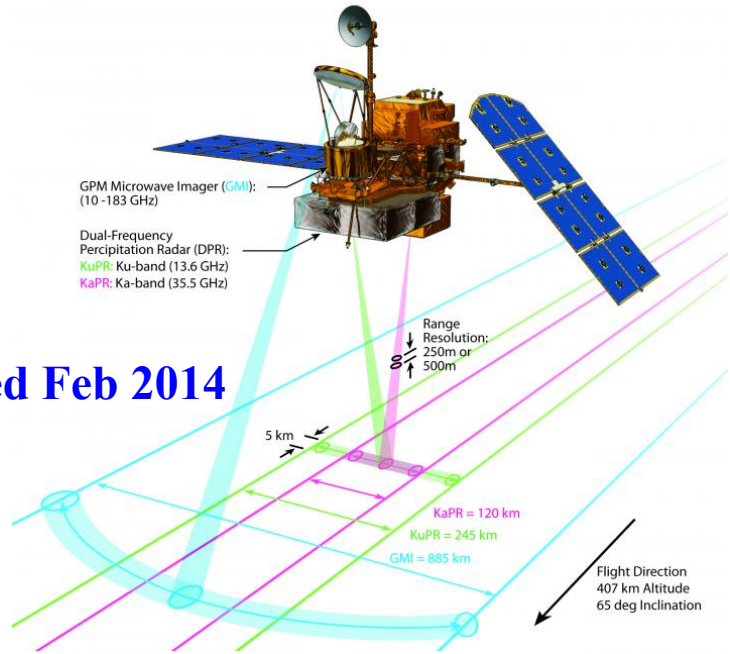
Simulations were carried out in NASA NAS Pleiades.

ESTO Earth Science Technology Office

Launched in 1997-2015

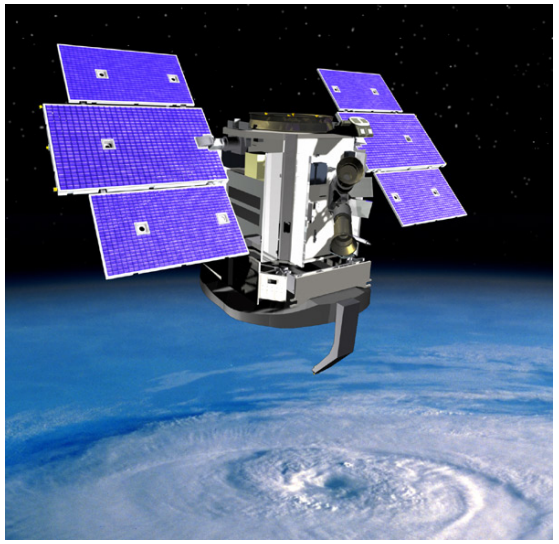


GPM



Launched Feb 2014

CloudSat



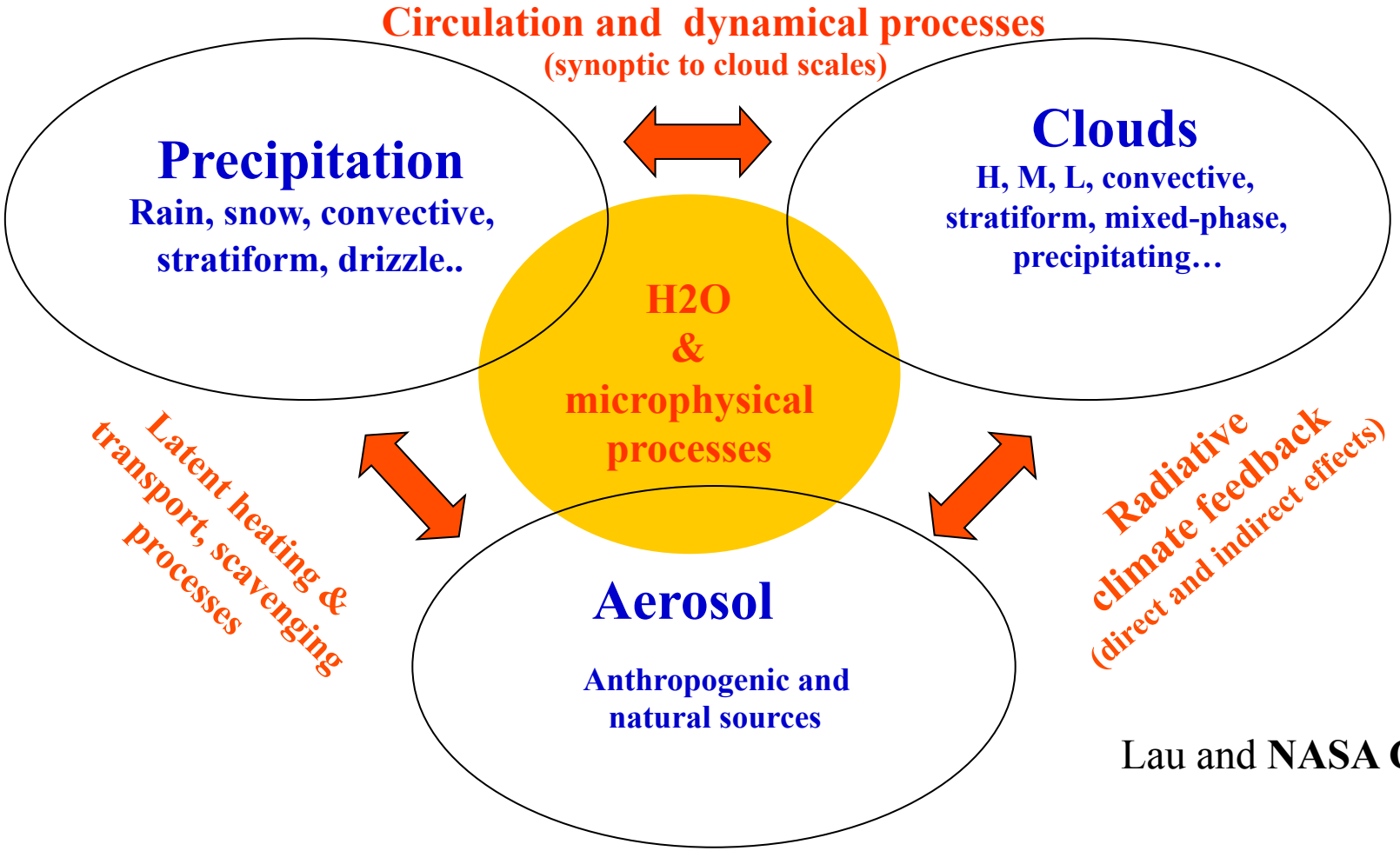
Launched
28 April
2006

CaPPM

+1

An Integrated Approach to Atmospheric Water Cycle and Climate Change Research

(satellite observations, field campaign, modeling, data processing and applications)



Lau and **NASA CaPMM**

Weather and climate models are using explicit microphysics schemes developed by CRM for their higher resolution forecast/simulation