EESI

FP7 Support Action - European Exascale Software Initiative

DG Information Society and the unit e-Infrastructures

EESI WP3 – Application Grand Challenges
Task 3.2: Working Group on WCES
Weather, Climatology and Earth Sciences

Chair: ENES-CMCC (Giovanni Aloisio)

Vice-Chair: INGV (Massimo Cocco)

Speaker: prof. Giovanni Aloisio

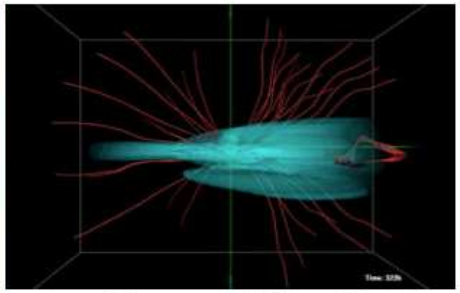
CMCC & University of Salento

Lecce - Italy

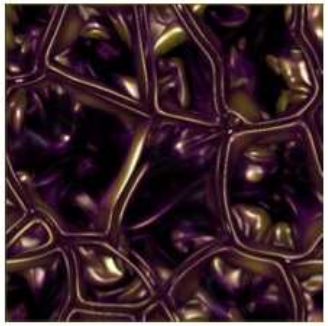
IESPl Meeting, San Francisco April 6, April 2011

Geoscience really does need Petascale Computing (and Beyond)

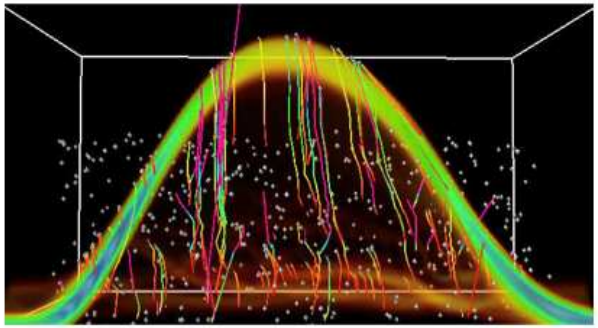
In a wide variety of domains…



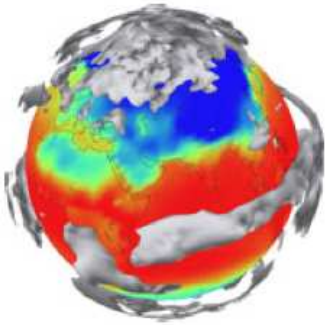
Space Weather



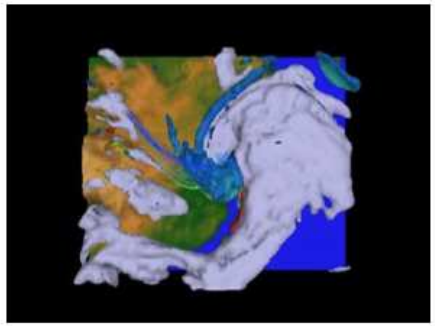
Turbulence



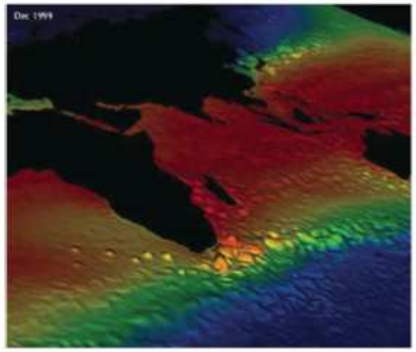
The Sun



Climate



Weather

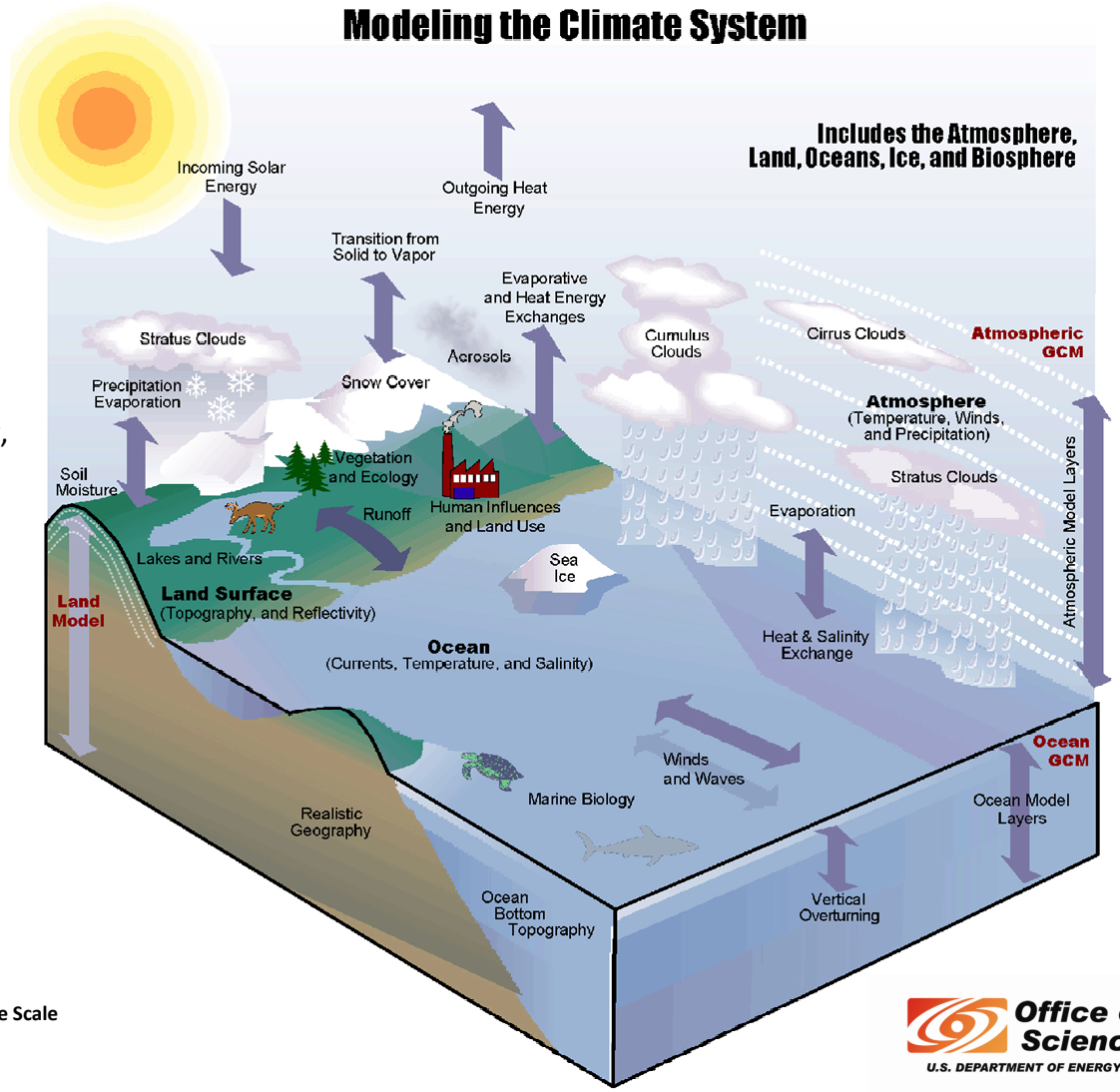


Oceanography

Climate Modeling

The big challenge is to model this complex system

- Several complex process to be simulated
- Several interacting processes
- Great range of time scales to be analyzed
- Great range of spatial scales to be considered
- Need interdisciplinar sciences (physics, chemistry, biology, geology,…)
- Inherently non-linear governing equations
- Need sophisticated numerics
- Need huge computational resources

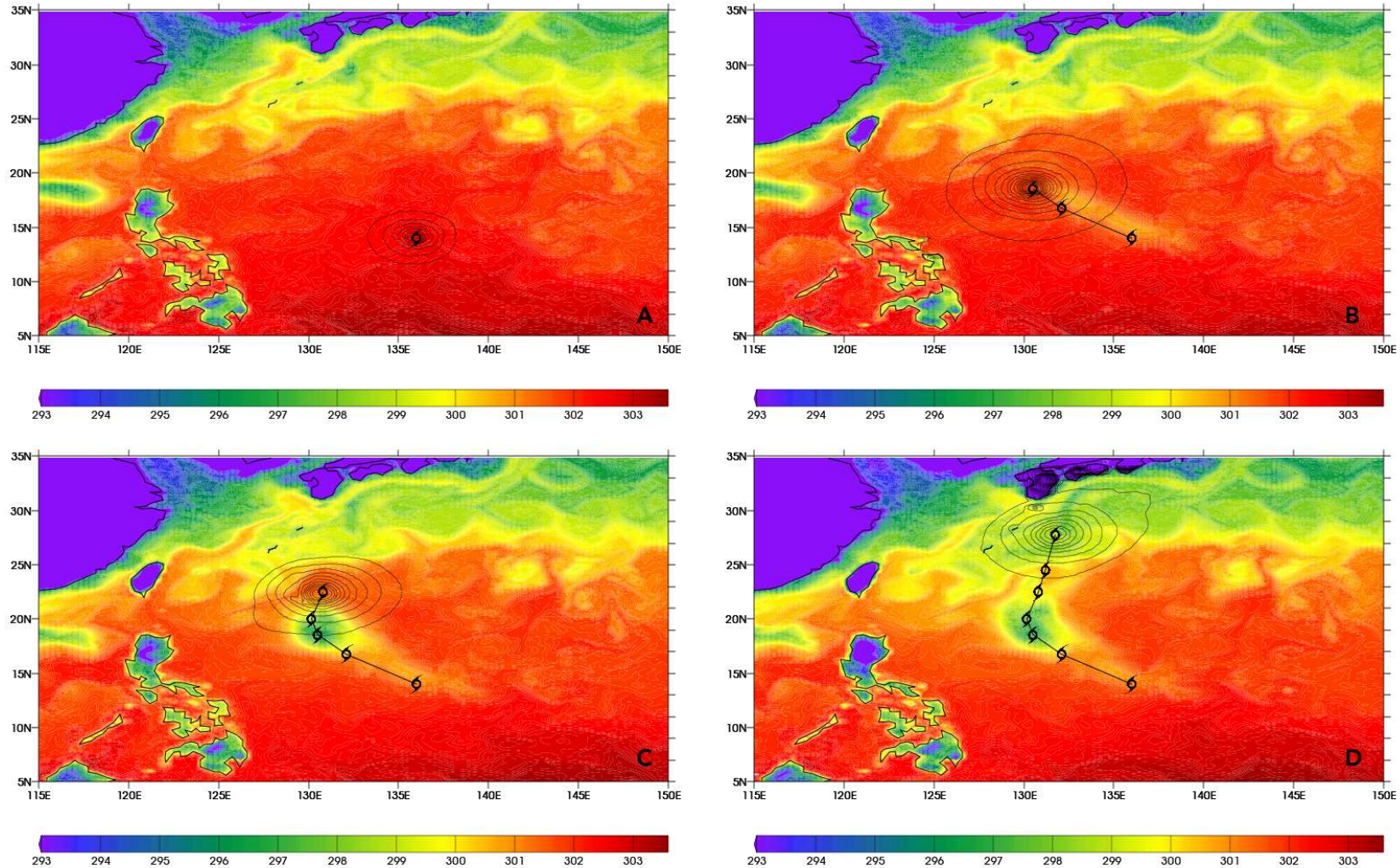


Warren M. Washington – NCAR

Scientific Grand Challenges Workshop Series:
Challenges in Climate Change Science and the Role of Computing at the Extreme Scale
DOE Workshop (ASCR-BER)
November 6-7, 2008

Office of Science
U.S. DEPARTMENT OF ENERGY

Modelers work on **Global Models of each climate component** to capture **small-scale features** (such as tropical cyclones) and **simulate how they interact with other components** (such as oceans)

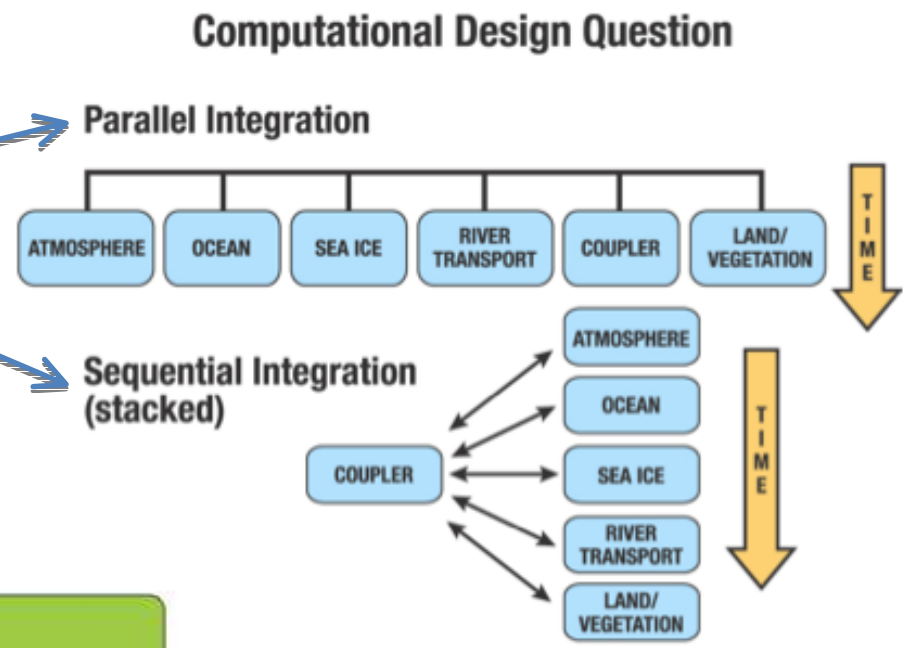


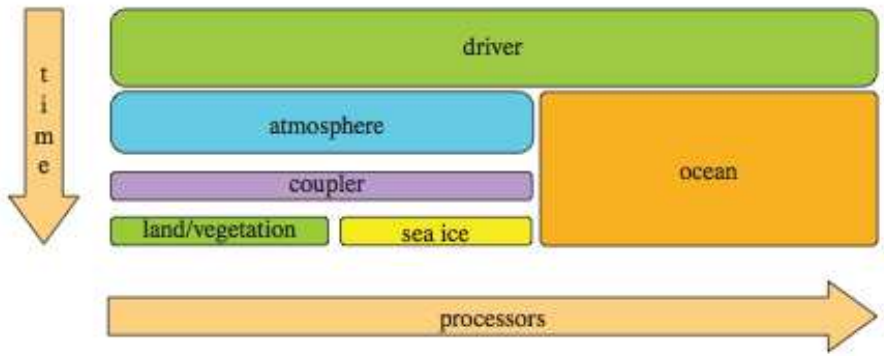
Warren M. Washington - NCAR

Coupling components in a climate model

Three methods of integrating coupled climate models

Hybrid





The flux coupler passes fluxes and variables from one component to another

W. Washington- Journal of Physics: Conference Series **16** (2005)

Global Atmospheric Models

The first developed (1960's) and evolved for application to the problems of

-deterministic weather forecasting → **N**umerical **W**eather **P**rediction (**NWP**)
-seasonal prediction
-climate simulations

NWP Models

Accurate prediction of the fluid flow
by applying the
highest resolution feasible

Climate Models

Extremely *long runs*
at
low resolution

Unified Models

The need to combine global and regional climate models is increasing

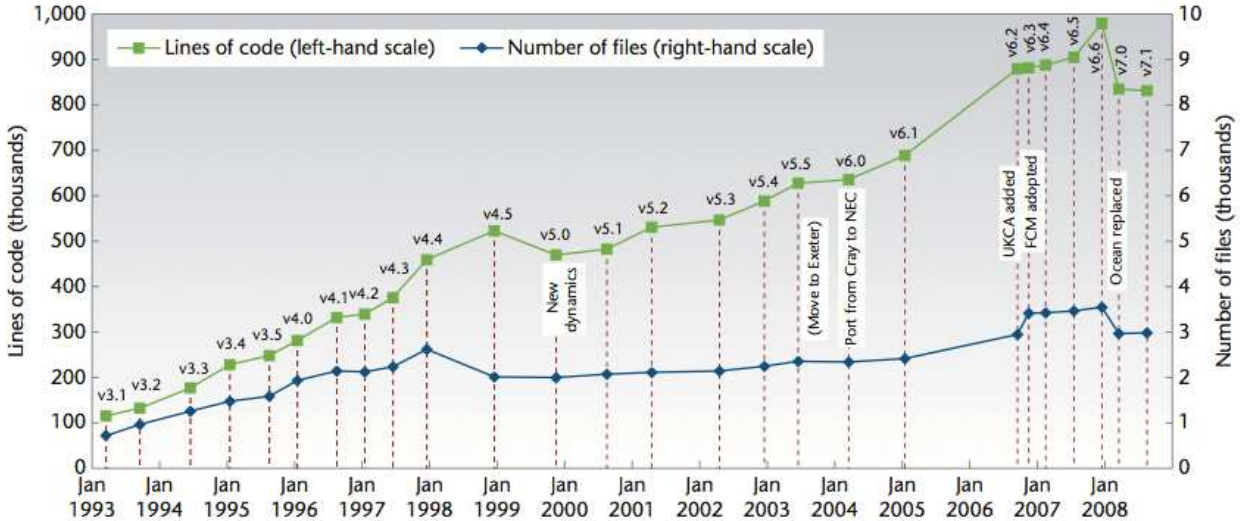


Figure 2. The Unified Model's growth over the past 15 years. Discontinuities in the growth curve indicate major component replacements, such as at versions 5.0 and 7.0, when the UM's dynamical core and ocean model were replaced, respectively. UM adopted the UK atmospheric chemistry model (UKCA) at version 6.2 and flexible configuration management (FCM) at version 6.3.

Growth over the past 15 years

The Met-Office Unified Model

Driver of change

1521-9615/09/$26.00 © 2009 IEEE
Copublished by the IEEE CS and the AIP

Steve M. Easterbrook
University of Toronto
Timothy C. Johns
Hadley Centre for Climate Prediction and Research

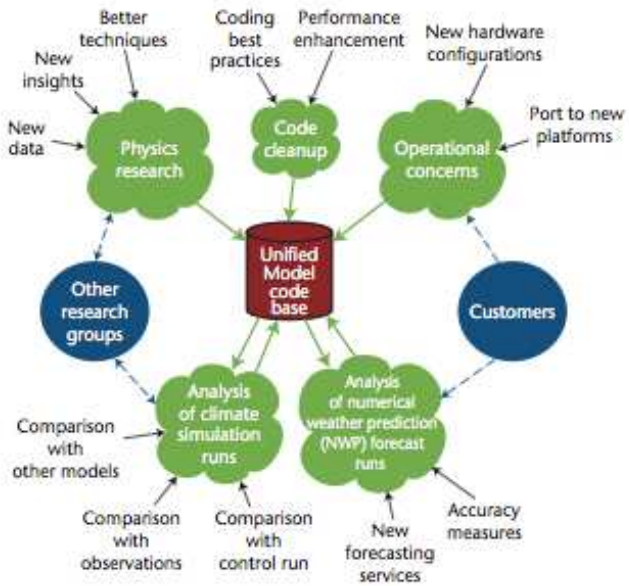


Figure 3. Drivers of change in UM code. The model's steady evolution is influenced by many factors, from advances in how scientists represent physical processes in the model to operational concerns, such as hardware upgrades. *(Courtesy of Damian Wilson, UK Meteorological Office.)*

The dream of future global climate models

Tornado simulation obtained by a **very high resolution "regional" atmospheric model** (WRF-Weather Research & Forecast Model)

WRF horizontal resolution

25 m

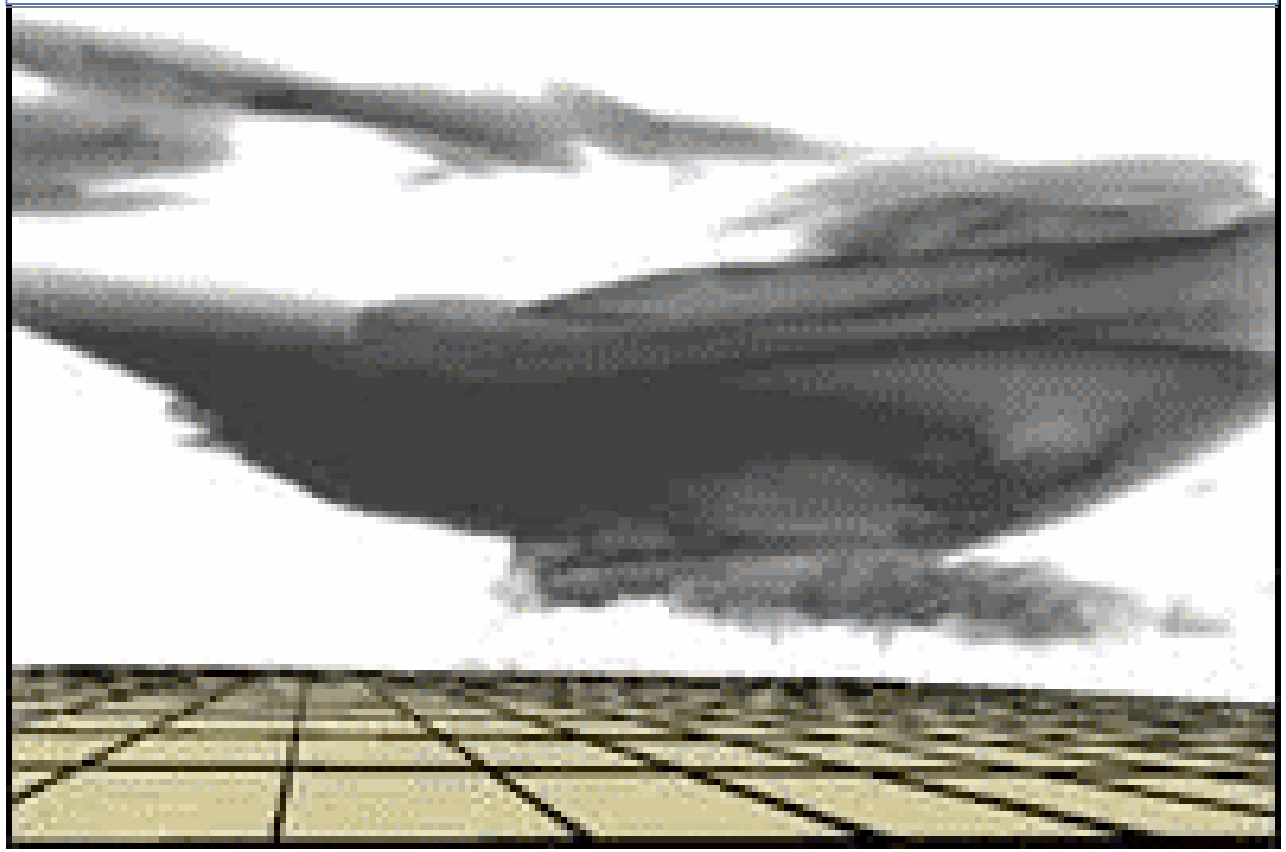
WRF vertical resolution

20 m

WRF domain size

Horizontal = 50 Km

Vertical = 18 Km



WRF number of grid cells

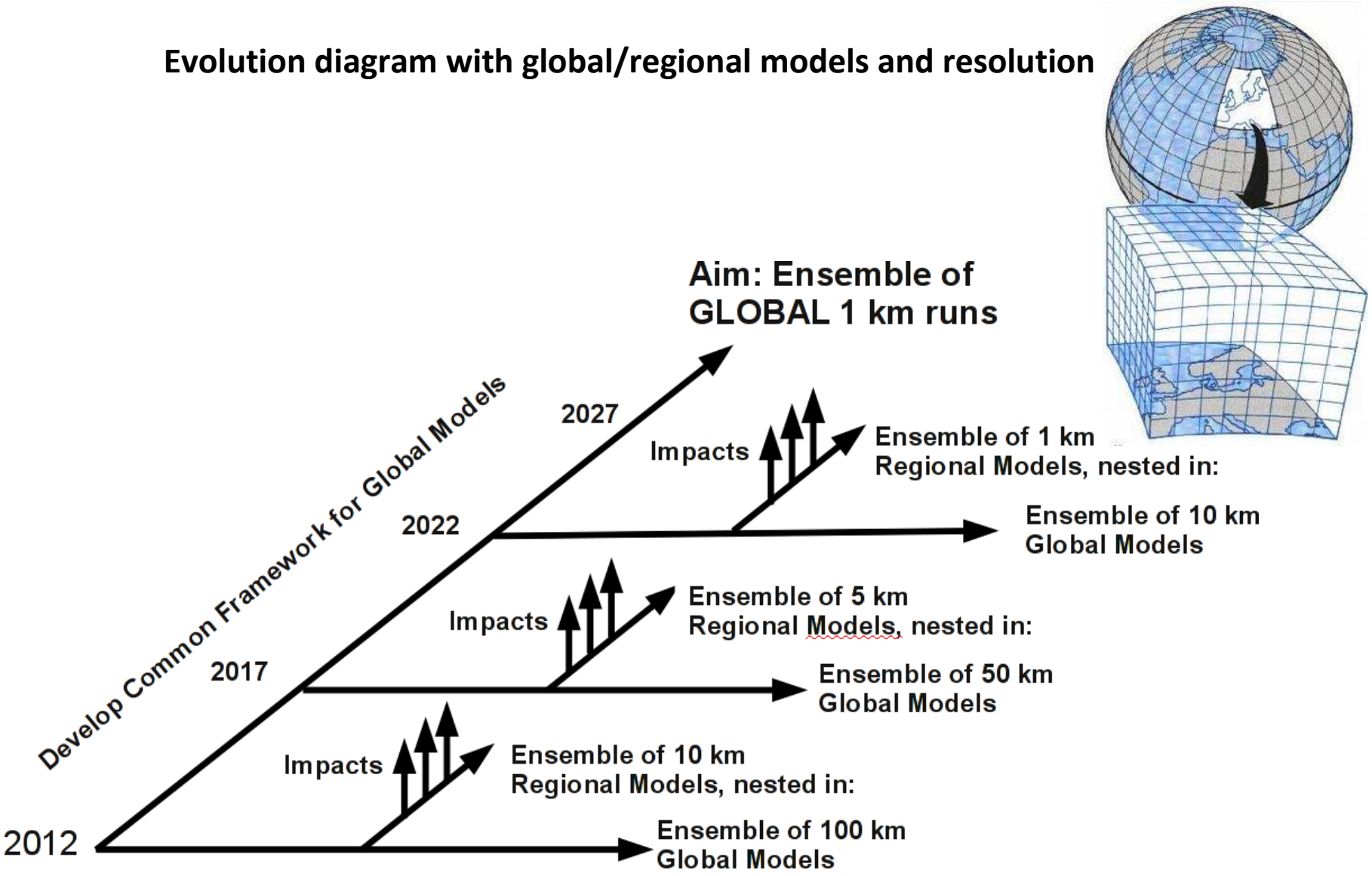
≈330 K

Computing Resources

Le Mieux Compurer
2048 processors
Pittsburgh
Supercomputing Center

The simulations was performed at the University of Pittsburgh (From Droegemeier and Xu of the University of Oklahoma's Center for Analysis and Prediction of Storms (CAPS))

Evolution diagram with global/regional models and resolution



Ensemble runs are needed for quantifying uncertainty

ENES foresight document (working draft)

Climate - basic equations

- Momentum conservation ('Newton's Second Law' → Navier-Stokes equations)

- Mass conservation ('Continuity equation')

Dynamics

- The ideal gas law: The thermodynamic state of the atmosphere, at any point, is determined by its pressure, temperature and density

- Energy conservation ('The first law of thermodynamics')

(Conservation of water (vapour, liquid), salinity, …)

The atmosphere in "Primitive Equations"

Dynamics

$$\frac{\partial u}{\partial t} + \vec{V} \cdot \nabla u + \omega \frac{\partial u}{\partial p} - fv + \frac{\partial \phi}{\partial x} = F_X$$

$$\frac{\partial v}{\partial t} + \vec{V} \cdot \nabla v + \omega \frac{\partial v}{\partial p} + fu + \frac{\partial \phi}{\partial y} = F_y$$

$$\frac{\partial \phi}{\partial p} = -\alpha$$

Navier-Stokes equations

$$\nabla \cdot \vec{V} + \frac{\partial \omega}{\partial p} = 0$$

Continuity Equation

Dynamical Core

$$\frac{\partial T}{\partial t} + \vec{V} \cdot \nabla T + \omega \frac{\partial T}{\partial p} - \alpha\omega / Cp = Q / Cp$$

Thermodynamic Energy Equation

$$p\alpha = RT$$

Ideal Gas Law

$$\frac{\partial q}{\partial t} + \vec{V} \cdot \nabla q \cdot + \omega \frac{\partial q}{\partial p} = S_q$$

Moisture conservation

u, v, ω, T, α, Φ, and q

- In principle: possible to solve (#unknowns = #equations).
- In practice: analytical solutions not possible (e.g. non-linearity)
- (Various filtered forms of the equations of motion…)

A very complex natural system to be simulated

- Vast temporal (secs to B yrs) and spatial (1000s of km to microns) scales
- Highly nonlinear behavior

Newton's Law, applied to the atmosphere

$$\rho \mathbf{g} - \nabla p + \nu \nabla^2 \mathbf{u} = \rho \left(\frac{\partial}{\partial t} + \mathbf{u}.\nabla \right) \mathbf{u}$$

Unpacks into..

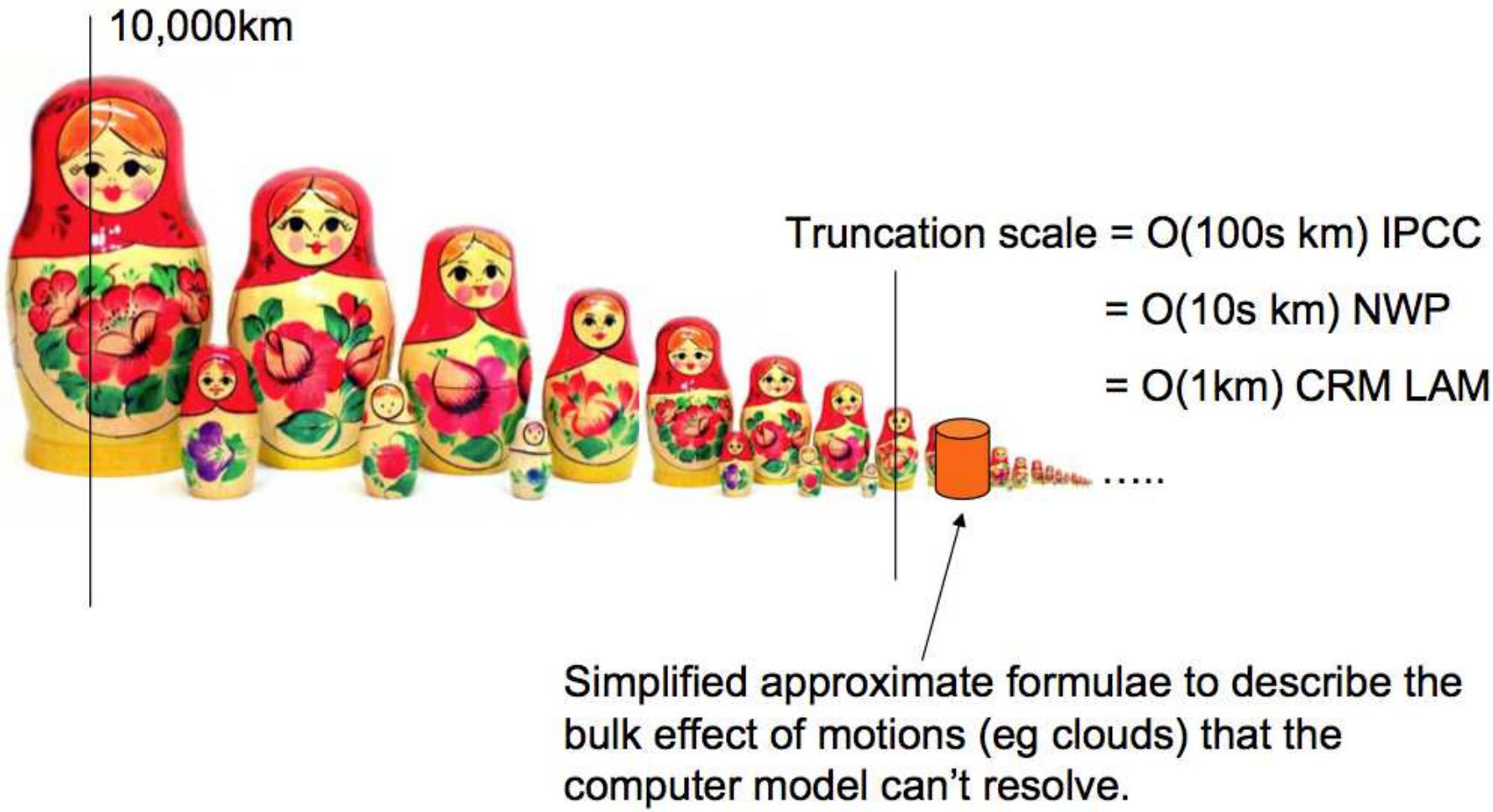
$$\rho \mathbf{g} - \nabla p + \nu \nabla^2 \mathbf{u} = \rho \left(\frac{\partial}{\partial t} + \mathbf{u}.\nabla \right) \mathbf{u}$$

← **nonlinear**

Unpacks into billions of individual equations, describing scales of motion from planetary scales to microscopic scales.

Tim Palmer - ECMWF

Even the world's biggest computer aren't big enough to represent all scales of motion in the atmosphere….



Tim Palmer - ECMWF

Dynamical cores issues at Exascale

- Dycores are the most challenging problem within geo-physical applications for future Exascale computing

  In today's ESM, *the dynamical cores are the main scalability bottlenecks*

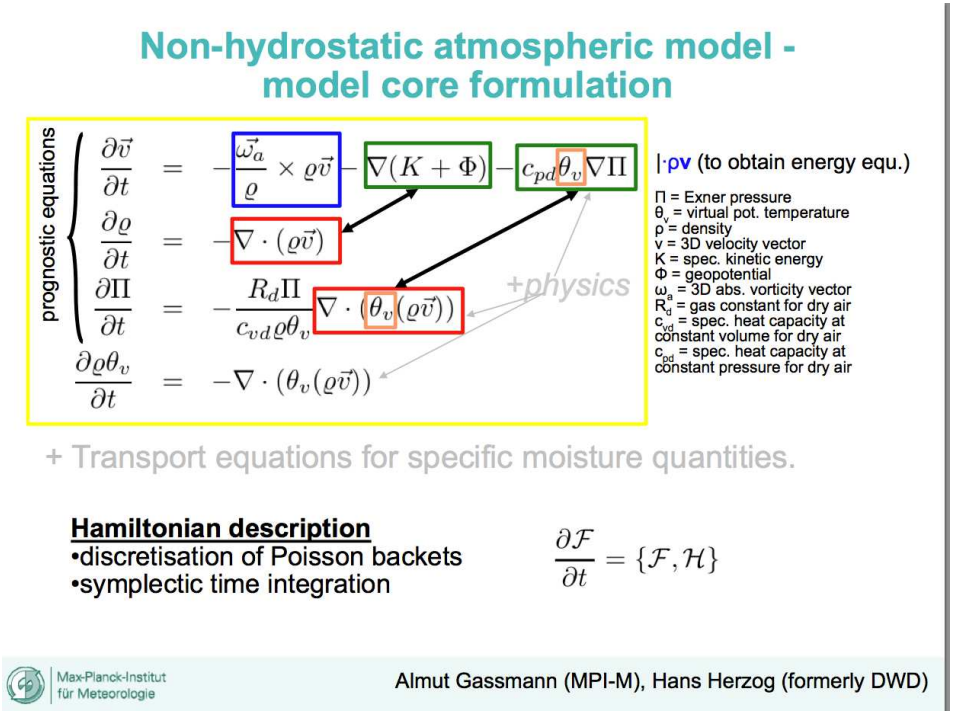
- The fundamental problem is the sequential nature of the time-stepping algorithm used to solve the PDEs involved

  As the resolution of the models increases, the time-step used normally needs to shorten in order to control errors caused by the time-discretization and to avoid instabilities.

Dynamical cores issues at Exascale

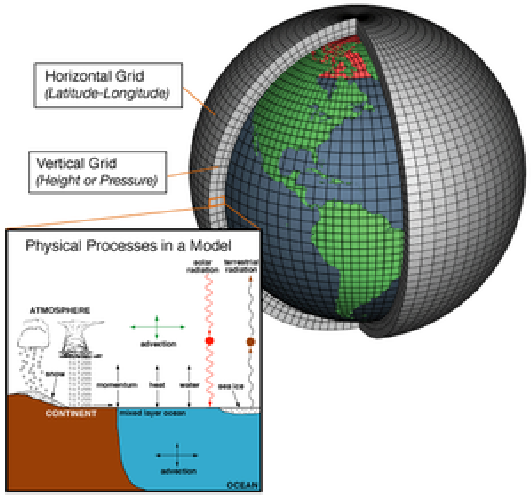
- In the timeframe of Exascale computing, we will enter the cloud-resolving scales for global atmospheric models (~1km)

This will necessitate moving to *non-hydrostatic dynamical cores*



Dynamical cores issues at Exascale

- The representation used for the physical quantities, like the *latitude longitude grid* seem to present special problems when going into the Exscale regime



The latitude longitude grid *suffers from the convergence* of the longitudes *when approaching the poles*. *More promising seem the various quasi-uniform grids* currently being developed to solve

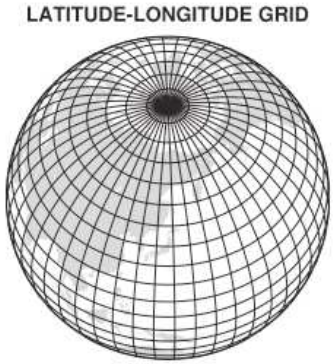
The Pole problem !

- Some currently used numerical scheme used to solve the Pole problem, like the spectral method (with its associated Gaussian grid)

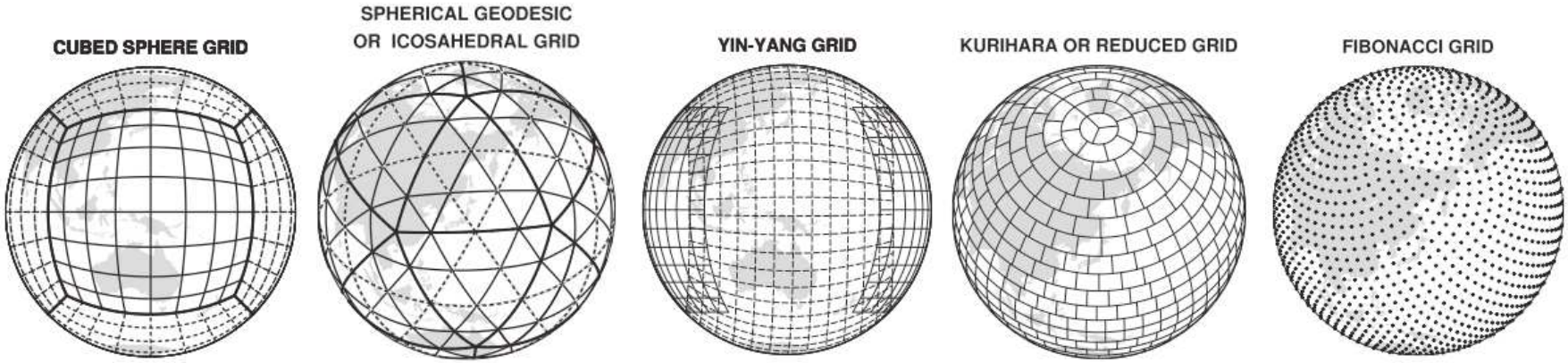
requires *global communications* (within the transforms between the spectral representation and the Gaussian grid) and may *scale poorly at very large core counts*

The Pole Problem

- The latitude longitude grid suffers from the convergence of the longitudes when approaching the poles



- Promising approaches are based on quasi-uniform grids (cubed sphere, icosahedral, Yin-Yang, Fibonacci …)

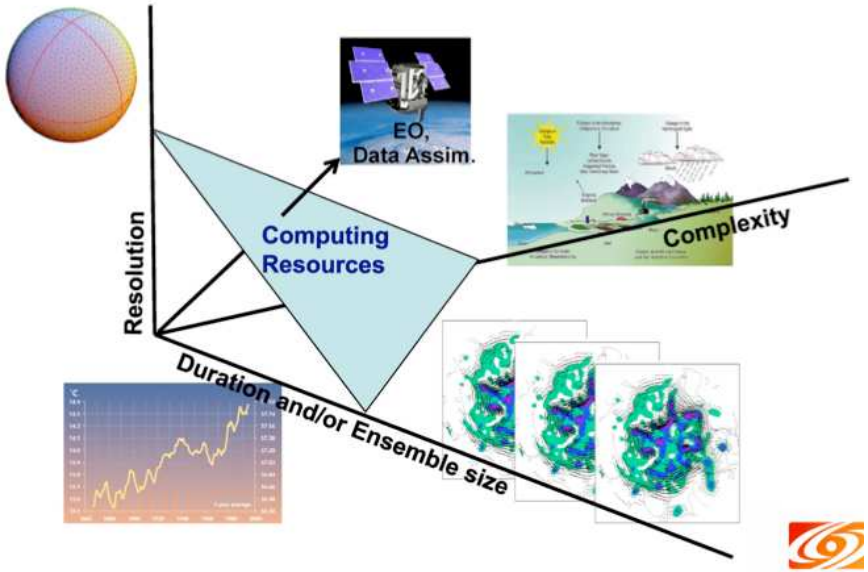


The 5 HPC dimensions of climate prediction

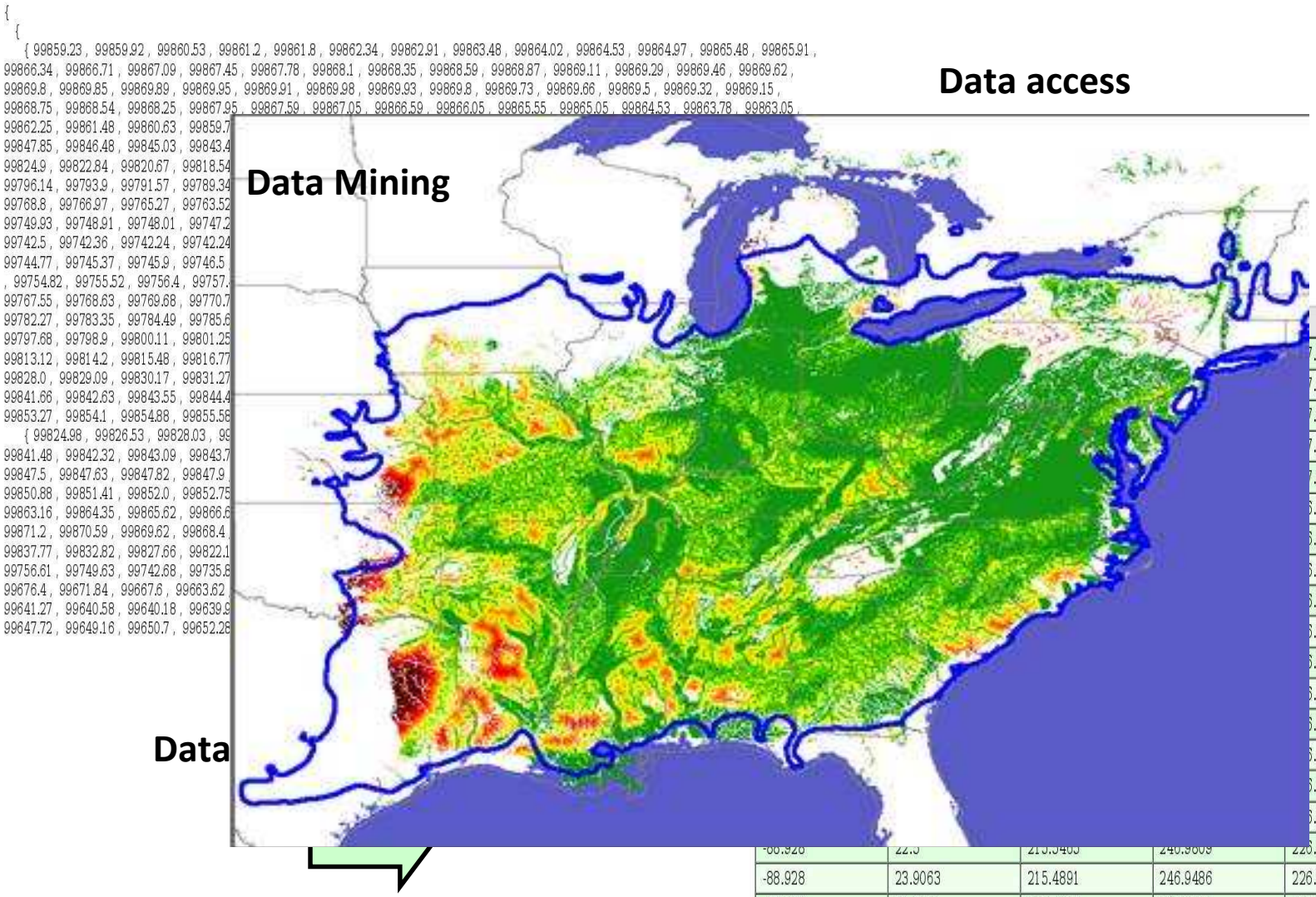
- Resolution
- Earth-System complexity
- Duration and/or Ensemble size
- Paleo-timescale integration
- Data assimilation and initial value forecasts (seamless prediction)

Capacity

Capability



Data access, analysis and mining



Average	Variance
.27002589	118.20635699
.20344162	118.71782431
.14841461	118.93681027
.08705139	119.32944630
.03001658	119.66761535
.96573385	120.08689023
.90545019	120.45400045
.84642792	120.79659695
.78913752	121.12353909
.72494253	121.61722149
.66145833	122.09671745
.59991582	122.48782104
.54492569	122.75315922
.48511887	123.20487250
.42746862	123.56921073
.37406540	123.81541520
.31255722	124.28107138
226.25874201	124.54055867
226.20121765	124.90133119
226.14244207	125.34587587

-88.928	23.9063	215.4891	246.9486	
-88.928	25.3125	215.4318	246.9181	
-88.928	26.7188	215.3580	246.8955	

FP7 Support Action - European Exascale Software Initiative

DG Information Society and the unit e-Infrastructures

EESI - WP 3 - Application Grand Challenges
Task 3.2: Working Group on WCES
Weather, Climatology and Earth Sciences

WG 3.2 Roadmap

EESI - WG 3.2	WCES (Weather, Climatology and Earth Sciences)			
Name	Organization	Email	Country	Area of Expertise
Giovanni Aloisio	ENES-CMCC	giovanni.aloisio@unisalento.it	IT	Exascale Computing
Massimo Cocco	INGV	massimo.cocco@ingv.it	IT	Seismology
Nadia Pinardi	CMCC	n.pinardi@sincem.unibo.it	IT	Oceanography
Sylvie Joussaume	IPSL	sylvie.joussaume@lsce.ipsl.fr	FR	Earth-system Modeling
Sophie Valcke	CERFACS	valcke@cerfacs.fr	FR	Coupled Climate Models
Jean-Pierre Vilotte	IPG	vilotte@ipgp.jussieu.fr	FR	Solid Earth, Seismology
Marie-Alice Foujouls	IPSL	Marie-Alice.Foujols@ipsl.jussieu.fr	FR	Climate Modeling
Bryan Lawrence	BADC	bryan.lawrence@stfc.ac.uk	UK	Climate Data Management
Graham Riley	Manchester Univ.	graham.riley@manchester.ac.uk	UK	High Performance Computing
Mats Hamrud	ECMWF	nar@ecmwf.int	UK	Data Assimilation
Tim Palmer	ECMWF	Tim.Palmer@ecmwf.int	UK	Weather/Climate Modeling
Reinhard Budich	MPI	reinhard.budich@zmaw.de	DE	Earth System Modeling
Joachim Biercamp	DKRZ	biercamp@dkrz.de	DE	High Performance Computing
Heiner Igel	LMU	heiner.igel@geophysik.uni-muenche	DE	Solid Earth, Geophysics
Colin Jones	SMHI	colin.jones@smhi.se	SE	Climate Change Modeling
Johan Silen	FMI	johan.silen@fmi.fi	FI	Geophysics & Computing
Jose Baldasano	BSC	jose.baldasano@bsc.es	ES	Earth Sciences

WG 3.2 Roadmap - Table of contents (I)

Table of contents

WG 3.2 Roadmap - Table of contents (II)

WG 3.2 Roadmap - Table of contents (III)

Critical Exascale Software for WCES

X-Stack Components	*Needed Capabilities*	*Exascale Uniqueness*	*Exascale Criticality*
Frameworks	Resilience API and utilities	3	C
	Multi-institutional/multiproject collaboration plan	2	U
	Tool chain development/selection	2	U
	Programming model evaluation/adoption	2	C
	Data placement	2	C
	Multicomponent simulation utilities	2	U
	Access to third-party libraries	1	C
Numerical Libraries	Fault-oblivious, error-tolerant software	3	C
	Asynchronous methods	2	C
	Overlap data and computation	3	U
	Self-adapting hybrid and hierarchical based algorithms	1	C
	Hybrid and hierarchical-based algorithms (e.g., linear algebra split across multicore and GPU)	1	U
	Algorithms that minimize communications	3	C
	Architecture-aware algorithms/libraries	3	C
	Autotuning-based software	1	U
	Standardization activities	1	U
	Energy-efficient algorithms	2	U
	Mixed arithmetic	1	U

Critical Exascale Software for WCES

X-Stack Components	Needed Capabilities	Exascale Uniqueness	Exascale Criticality
Algorithms	Scalability	2	N
	Fault tolerance/resilience	1	N
	Conforming to architectural requirements	3	N
	New areas/uses of algorithms	1	U
Debugging	Concurrency and architecture driven high frequency of errors/failures	3	C
	Scalability of debugger methodologies (data volumes and interfaces)	3	C
	Focus on multilevel debugging, communicating details of faults between software layers	3	U
	Synthesis of fault information into understanding in the context of apps and architecture	3	C
	Specialized lightweight operating systems	2	N
	Automatic triggers, need compile time bridge to debugging that removes need to rerun	2	N
	Scalable clustering of apps process states and contexts, filter/search within debugger	2	N
	Vertical integration of debug and per information across software layers	2	N
	Excision of buggy code snippets to run at lower concurrencies	1	N
	Heterogeneity	1	N

Critical Exascale Software WCES

X-Stack Components	Needed Capabilities	Exascale Uniqueness	Exascale Criticality
I/O	Customization with I/O, purpose-driven I/O	3	C
	New I/O models, SW, runtime systems and libs	3	C
	Intelligent/proactive caching mechanisms for I/O	3	N
	Fault-tolerant mechanisms	3	C
	I/O into programming models and languages	3	N
	Balanced architectures with newer devices	2	N
	File systems or alternative mechanisms	2	N
	Active storage	2	N
	Wide-area I/O and integration of external storage systems	2	N
	Special-purpose network protocols for parallelism	2	N
	Balanced architectures with newer devices embedded with the node	1	N
Scientific data management	Scalable data analysis and mining SW and tools	3	C
	Scalable data format and high-level libraries	3	C
	Scientific workflows tools	2	C
	Search and query tools	2	N
	Wide-area data access movement and query tools	2	N
	Scientific databases	2	N

Critical Exascale Software WCES

X-Stack Components	*Needed Capabilities*	*Exascale Uniqueness*	*Exascale Criticality*
Programming models	Exascale programming model	3	C
	Scalable, fault-tolerant MPI	3	C
	Applications development tools	3	N
	Heterogeneous node programming model	2	C
	Domain-specific programming models	2	N
	Language features for massively parallel I/O	2	U
	Language support for adaptive computation	2	N
	Interoperability between models	1	2N
Compilers	Implement exascale languages	3	C
	Support for resilience	3	C
	Implement heterogeneous programming models	2	C
	Support for massive I/O	2	C
	New optimization frameworks	2	N
	Interactions between compilers and tools, runtime	2	C
	Dynamic compilation, feedback optimization	2	N
	Autotuning-based software	2	N
	Enhancements to existing languages/APIs	1	N
	Automatic parallelization	1	N

Critical Exascale Software WCES

X-Stack Components	*Needed Capabilities*	*Exascale Uniqueness*	*Exascale Criticality*
Operating Systems	Define the base OS (Standard API)	3	C
	APIs for resilience (access to RAS, etc)	3	C
	Collective OS operations	3	N
	Scalable system simulation environment	2	C
	Improved APIs for scalable performance monitoring and debugging	2	C
	New APIs for energy management	2	U
	Improved APIs for explicit memory management	1	C
	Improved APIs for threading	1	U
Performance	→ Extremely scalable performance methods and tools	3	C
	→ Performance measurement and modeling in presence of noise/faults	3	C
	→ Automated/automatic diagnosis and autotuning	2	N
	Predictive future large-scale system design	2	C
	Vertical integration across SW layers	2	N
	→ Performance-aware design and implementation	2	U
	Performance optimization for other metrics than time	2	U
	→ Support for heterogeneous hardware and hybrid programming models	1	C

X-Stack Components	Needed Capabilities	Exascale Uniqueness	Exascale Criticality
Programmability	New models of computation	3	C
	New runtime/OS interface for environment aware programming	2	C
	Programmability to decouple exascale system issues from applications programming	1	C
Power	Power performance monitoring and aggregation that scales to 1 billion core system	3	C
	Power control system	3	C
	Scalable control algorithms to bridge gap between global and local power models	2	C
	Power-aware and scalable resource control and scheduling	2	C
	Optimally tuned system power based on control loop	1	N
	Power instrumentation and control standardization	1	N
Runtime Systems	Load balance	3	C
	Asynchrony, overlap	2	C
	Hierarchical execution models and scheduling	3	N
	Scaling/optimization of communications	3	C
	Memory management and locality scheduling	2	C
	Heterogeneity: scheduling	2	U
	Fine-grained mechanisms at node level	1	U

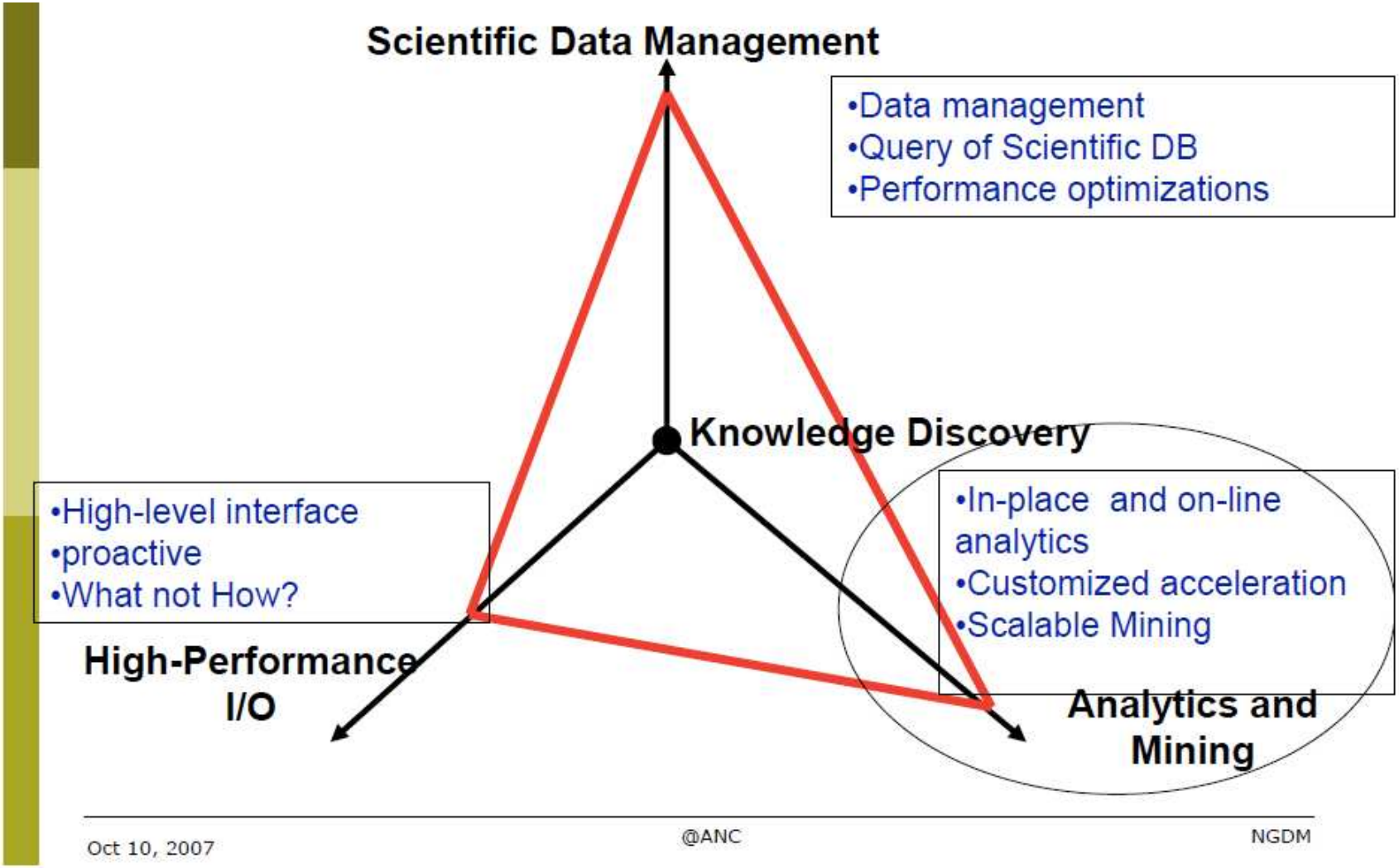
Critical Exascale Software

X-Stack Components	Needed Capabilities	Exascale Uniqueness	Exascale Criticality
	Performance measurement and modeling in presence faults	3	C
	Better fault tolerance protocols	2	C
	Fault isolation/confinement	2	C
	NV-RAM for local state, cache of file system	2	C
	Replication (TMR, backup core)	2	U
	Proactive actions (migration)	2	U
	Domain-specific API and utilities for frameworks	2	C
	Applications-guided fault management	2	C
	Language/compiler/runtime support for resilience (fault-aware programming, API from OS, RAS)	2	C
Resilience	Fault-tolerant MPI	2	C
→	Fault-oblivious, error-tolerant numerical libraries	2	C
	Resilient applications and algorithms	1	N
→	Fault-oblivious system software	2	C
	Fault-aware system software and API for resilience	2	C
→	Prediction for time optimal checkpoint/migration	2	U
	Fault models, event log standardization root cause analysis	2	C
→	Resilient I/O, storage, and file systems	2	C
	Situational awareness	2	C
	Experimental environment	2	C
	Fault isolation/confinement + local management	2	C

Critical Exascale Software for WCES

X-Stack Components	Needed Capabilities
Numerical libraries	Fault-oblivious, error-tolerant software, Asynchronous methods, Overlap data and computation, Hybrid and hierarchical based algorithms, Algorithms that minimize communications
Algorithms	Scalability, Fault-tolerance/resilience, New areas/uses of algorithms
Debugging	Scalability of debugger methodologies, Focus on multilevel debugging, Synthesis of fault information into undestanding in the context of apps, Heterogeneity
I/O	New I/O models, SW, run time systems and libs, Intelligent/proactive caching mechanism for I/O, Fault-tolerant mechanisms, Active storage, Wide-area I/O and integration of external storage systems
Scientific data management	Scalable data analysis and mining SW and tools, Scalable data format and high-level libraries, Search and query tools, Wide-area data access movement and query tools, Scientific databases
Programming models	Exascale programming models, Heterogeneous node programming model
Compilers	Implement heterogeneous programming models, Support for massive I/O, Auto-tuning based software, Automatic parallelization
Performance	Extremely scalable performance methods and tools, Performance measurement and modeling in presence of noise/faults, Automatic diagnosis and auto-tuning, Peformance-aware design and implementation, Support for heterogeneous hardware and hybrid programming models
Power	Power performance monitoring and aggregation, Power control system, Power-aware and scalable resource control and scheduling, Optimally tuned system power based on control loop
Runtime system	Load balance, Asynchrony, overlap, Optimization of communications, Memory management and locality scheduling
Resilience	Fault-oblivious system software, Prediction for time optimal checkpoint, Resilient I/O, storage and file systems

Challenges in Scientific Knowledge discovery



Alok Choudary

Knowledge discovery from climate change data

Main challenge: large scale, scalable and efficient scientific knowledge discovery from data

Climate data: multi dimensional, multi model, multi resolution,…

Mining ensembles (new dimension!) to “face” uncertainty

Multiple challenges at different levels:

- Efficienti data indexing
- Multi-layer parallel I/O design and implementation
- MPI-based analytics functions/libraries
- MPI-based mining functions/libraries
- Active storage primitives (move analytics and mining functions to the storage nodes) & parallel file systems

CMCC Research Activity on Scientific Data Management and possible contribution to the X-Stack software

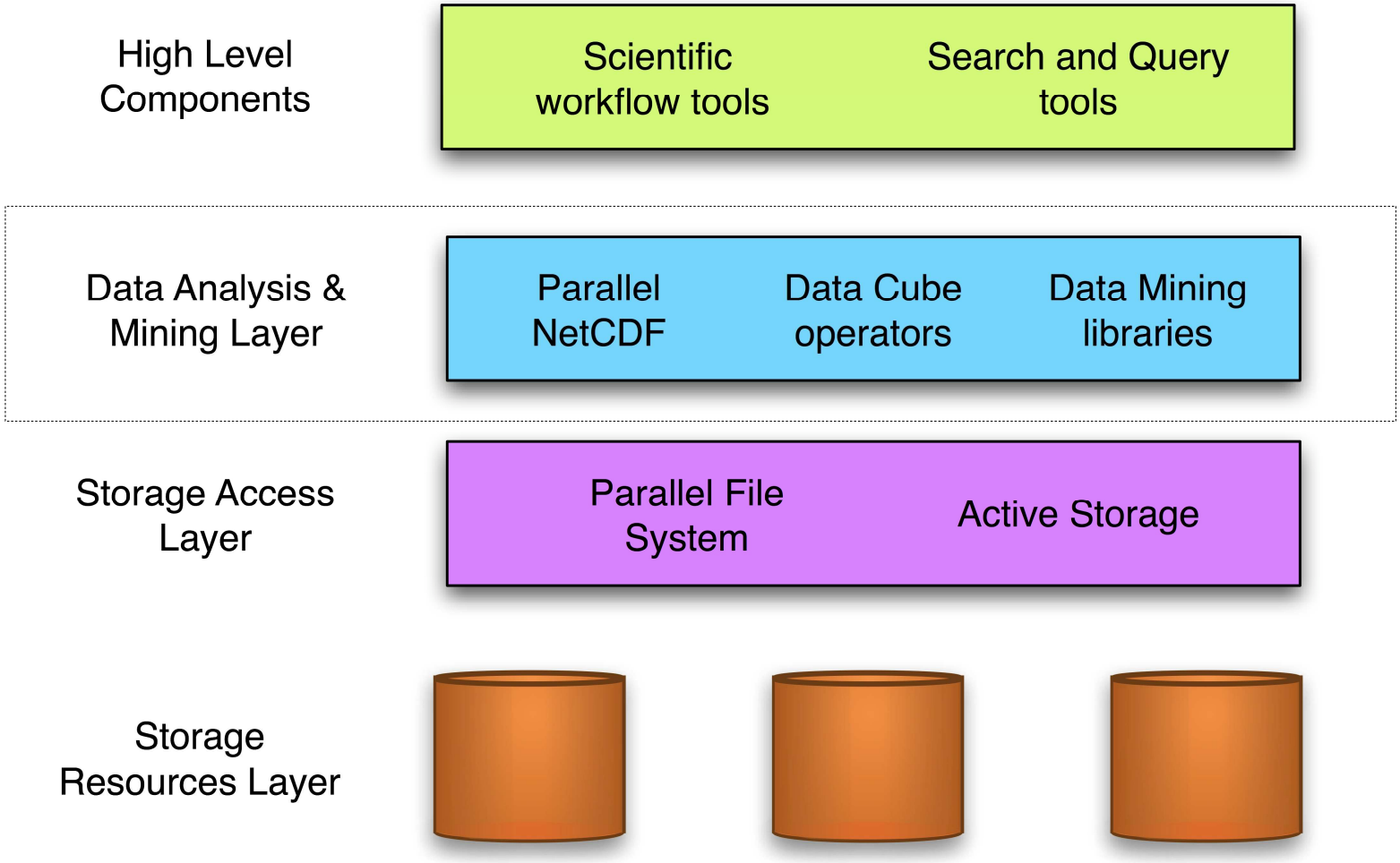
Strong interest in Scientific data management activities:

Scientific data management	Scalable data analysis and mining SW and tools
	Scalable data format and high-level libraries
	Scientific workflows tools
	Search and query tools
	Wide-area data access movement and query tools
	Scientific databases

and some I/O related topics:

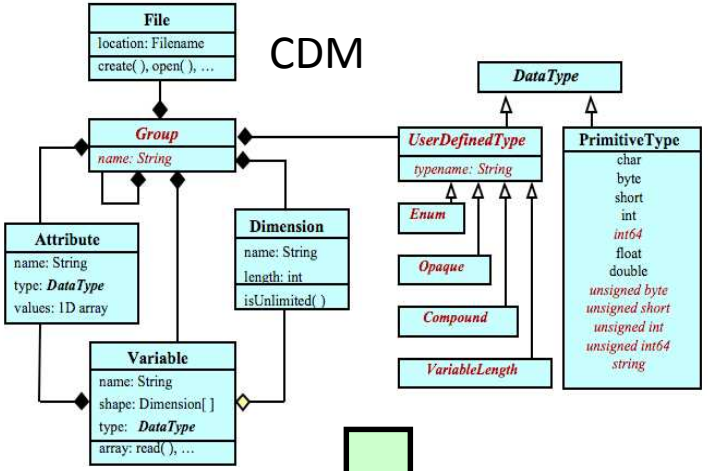
I/O	Active storage
	Wide-area I/O and integration of external storage systems

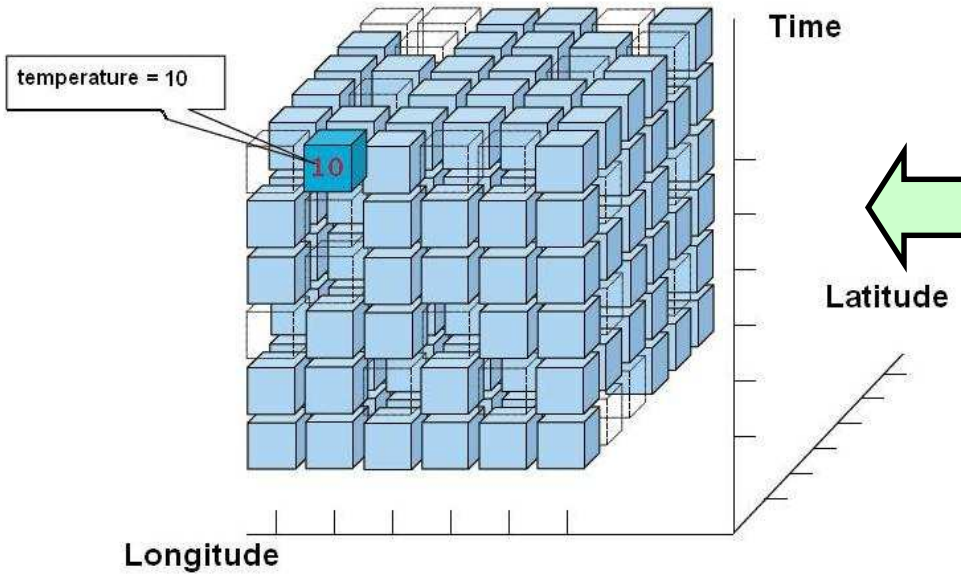
CMCC Research Activity on Scientific Data Management and possible contribution to the X-Stack software



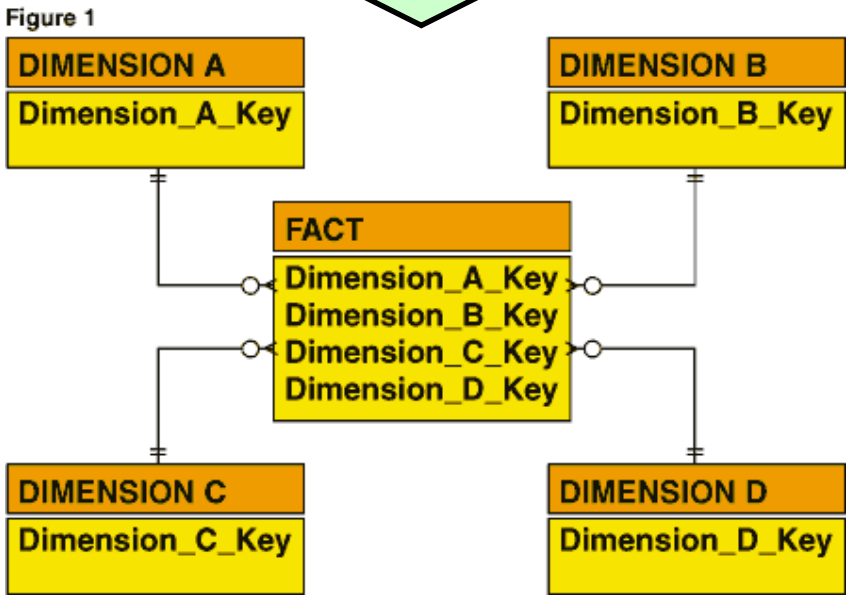
From NetCDF data to data cube virtualization

Data Formats



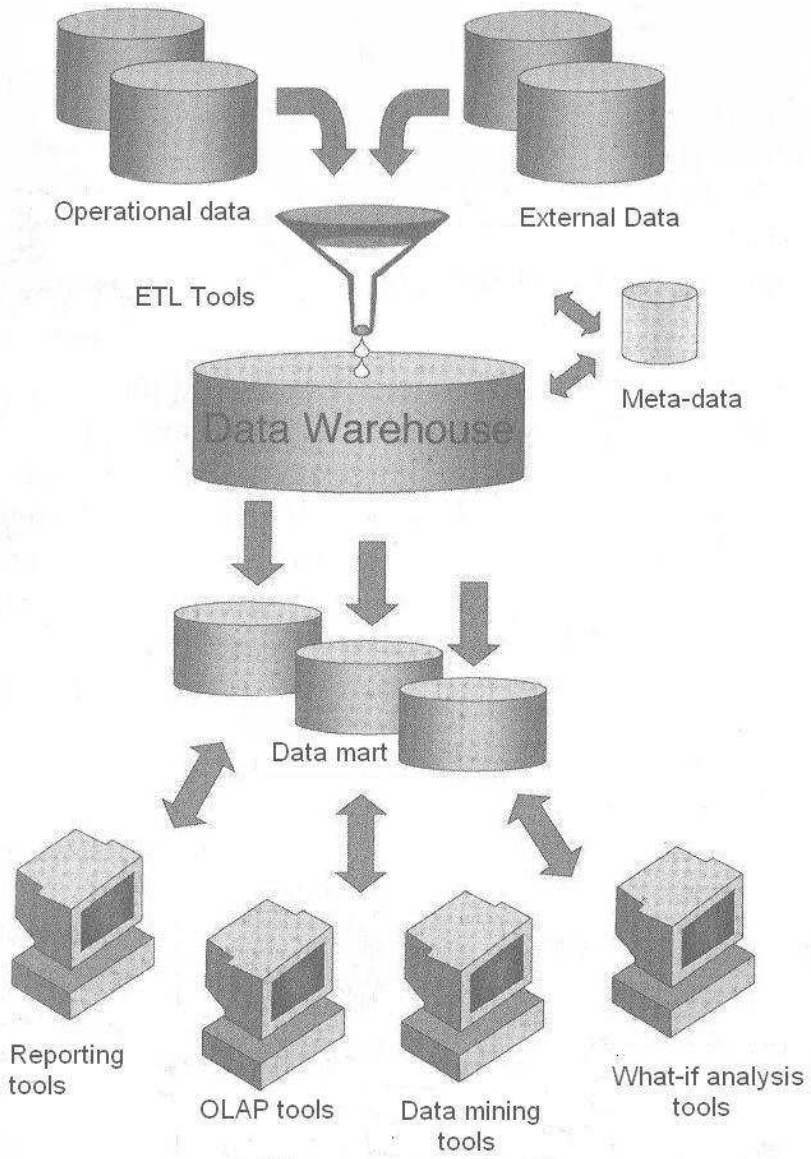


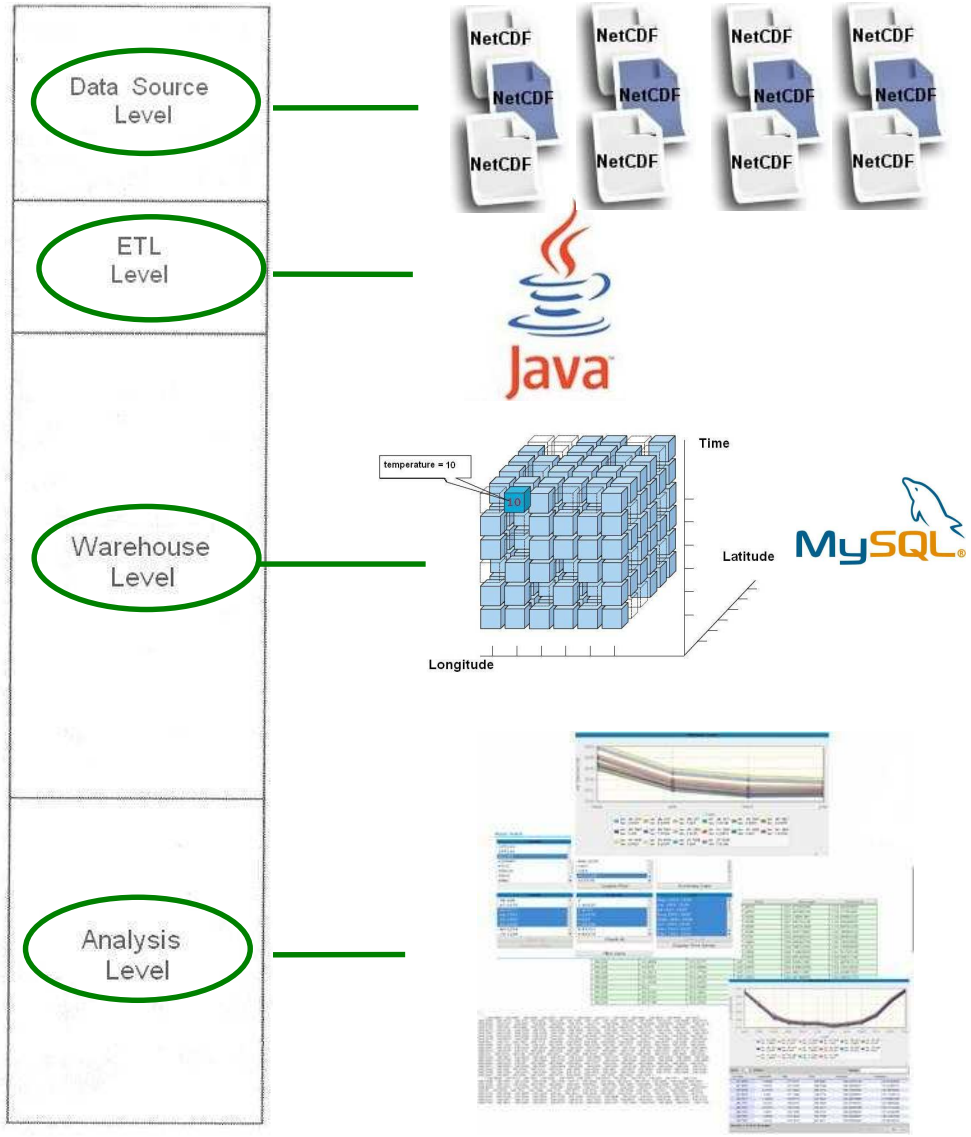
Data cube implementation



Star Schema

Extraction, transformation & Loading





Thanks