HPC and the progress of weather and climate forecasting

Philippe Bougeault

Météo-France

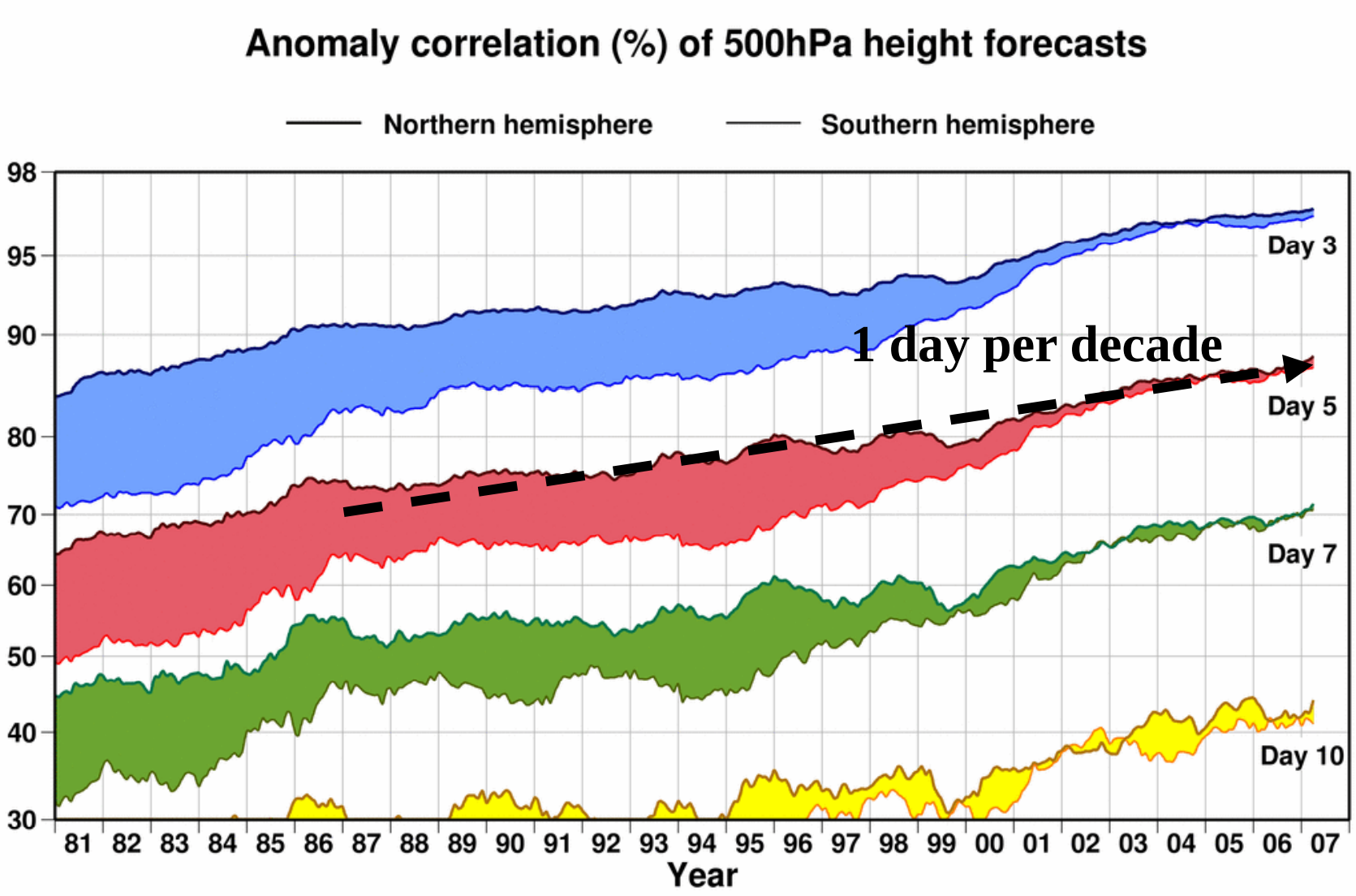
and

European Center for Medium-range Weather Forecasts

Outline

- **Examples of progress of weather forecasting over the last 20 years**
- **Where did the progress come from?**
- **Some examples of optimizations**
- **The Climate Problem: towards a convergence between climate and weather models?**

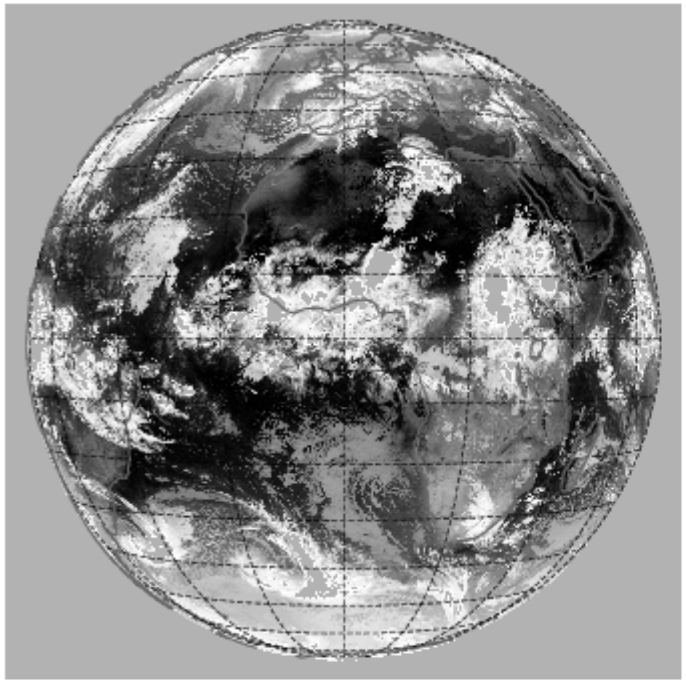
Improvements of weather forecasts skill over the last 30 years



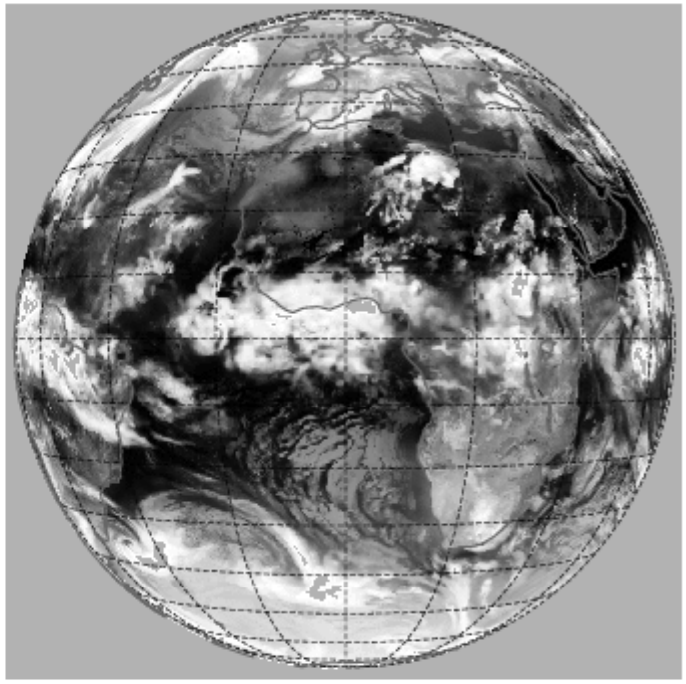
Simulated Meteosat imagery as a check of “model realism”

T799 36h forecast from 20080525

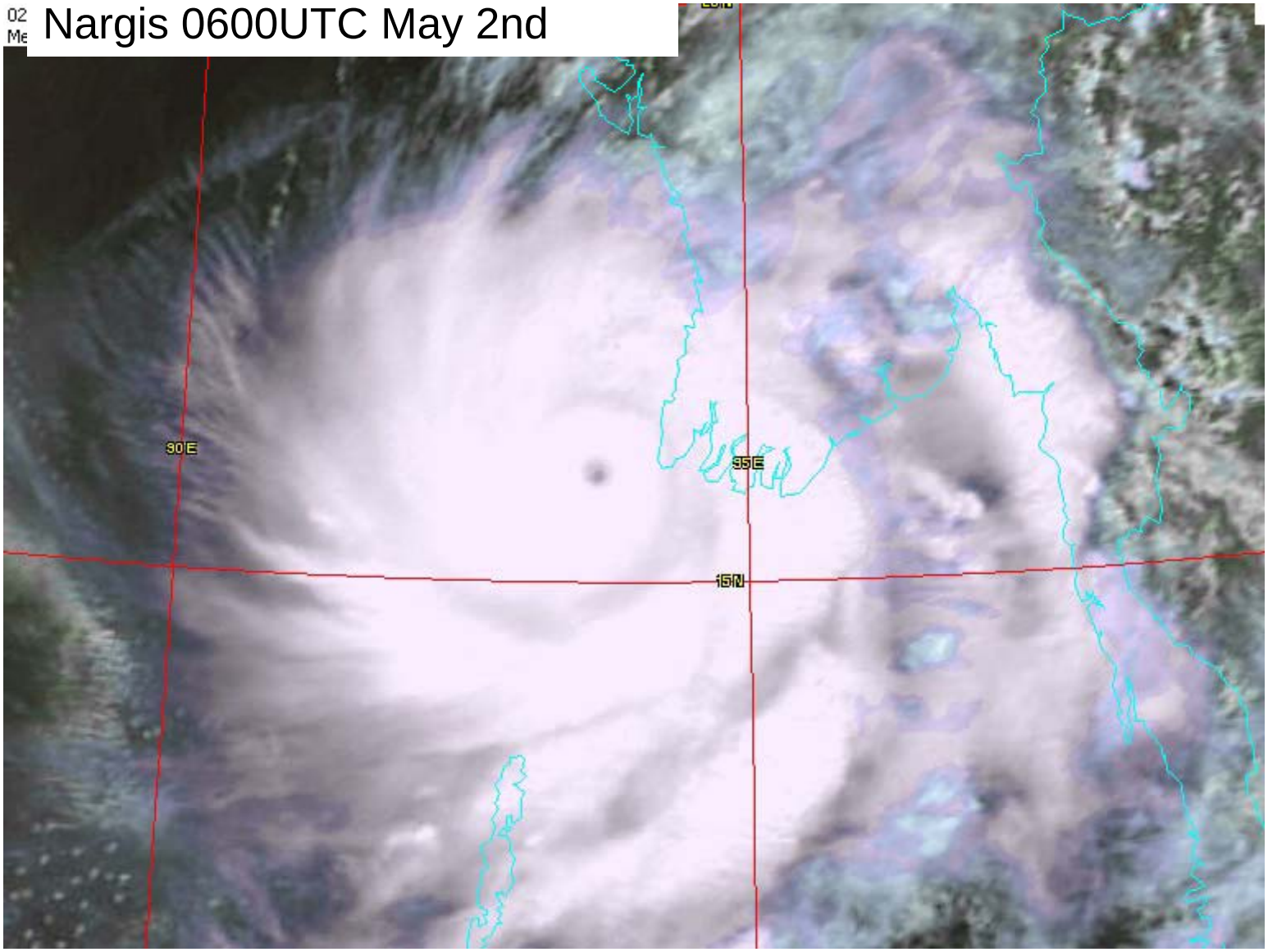
Meteosat 9 IR10.8 20080525 0 UTC



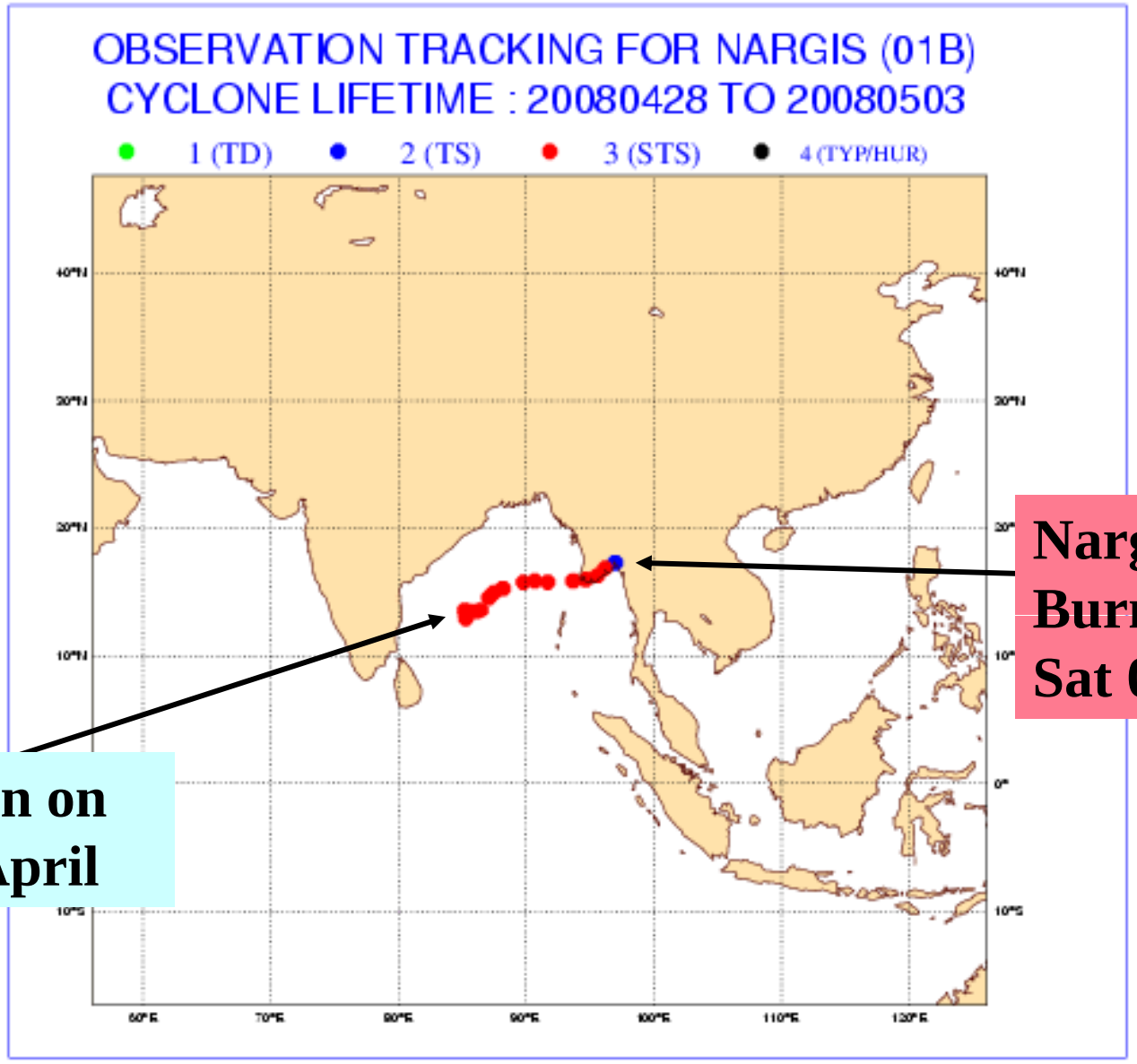
RTTOV gen. Meteosat 8 IR10.8 ECMWF Fc 20080525 00 UTC:



Exemple de la prévision de Nargis, mai 2008



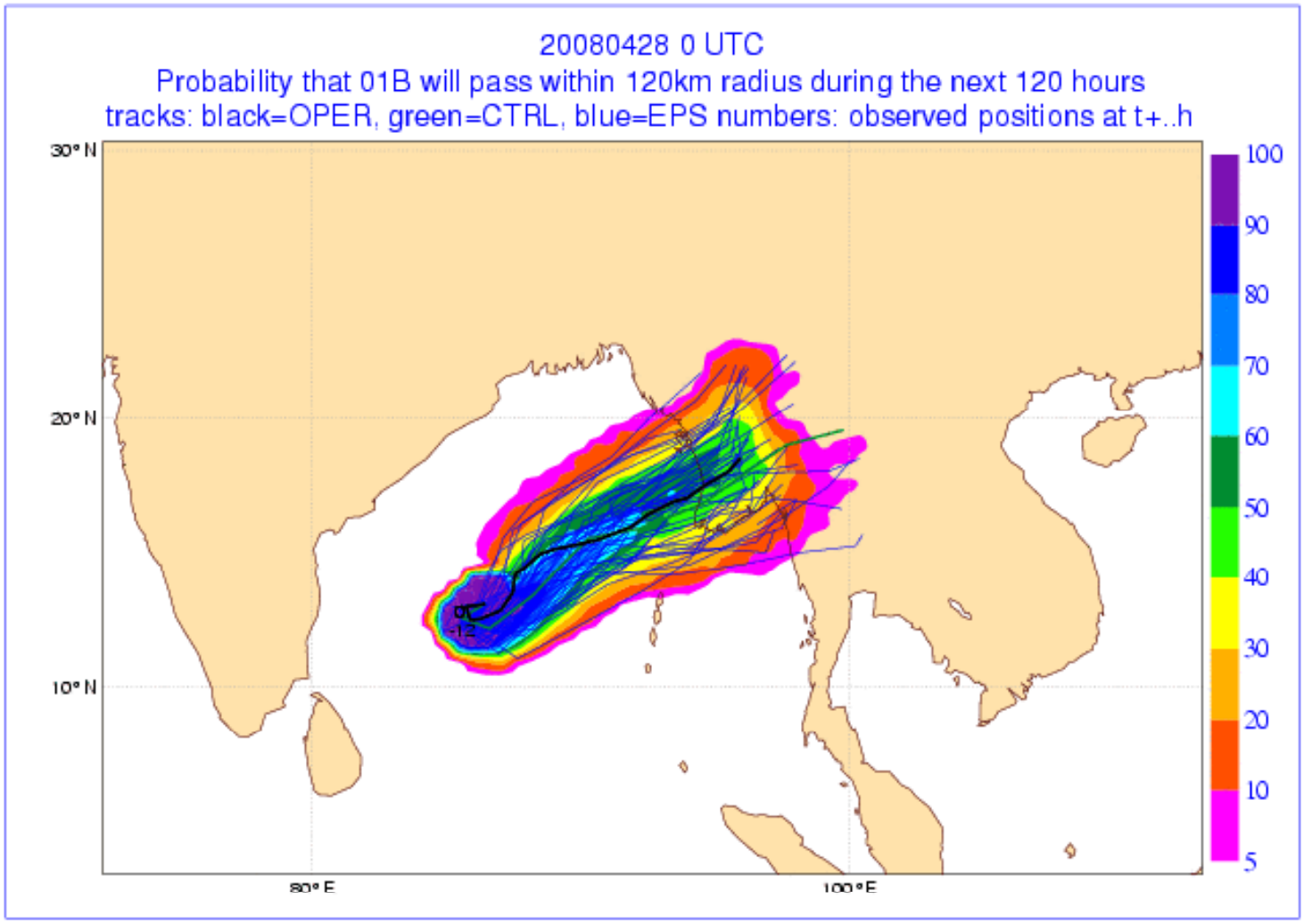
Tropical Cyclone Nargis – observed track



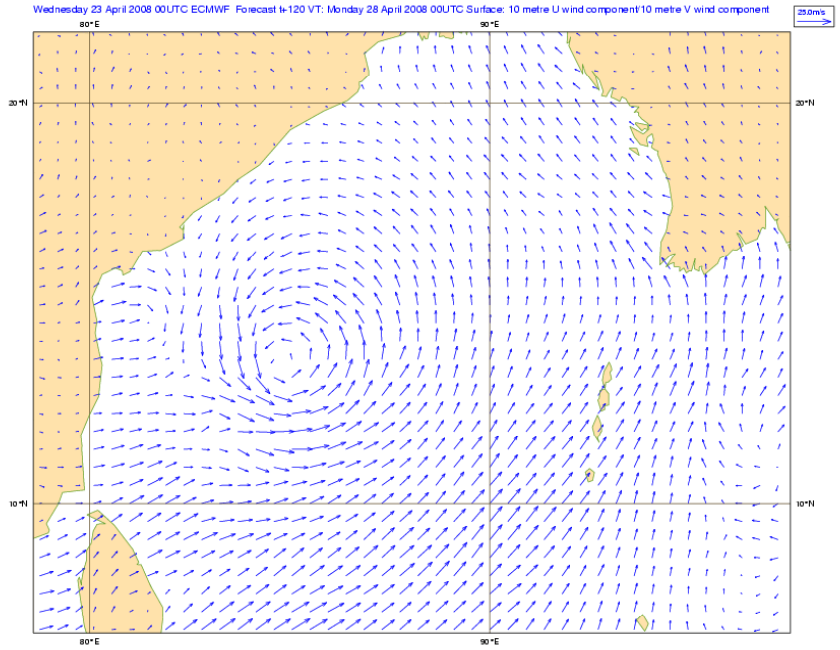
Nargis hits Burma on Sat 03 May

Formation on Mon 28 April

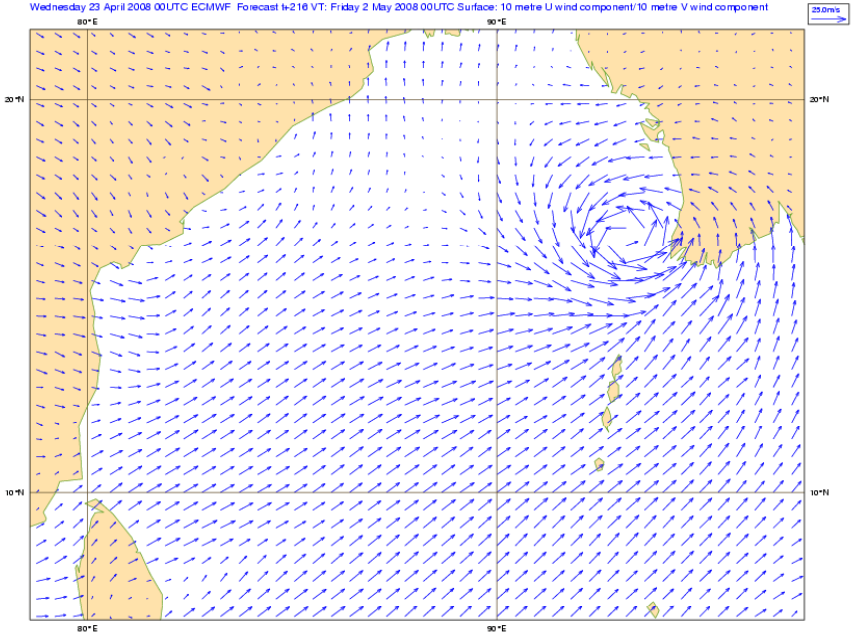
Ensemble forecasts from Monday 28 April 00Z: Accurate forecasts 5 days before the event



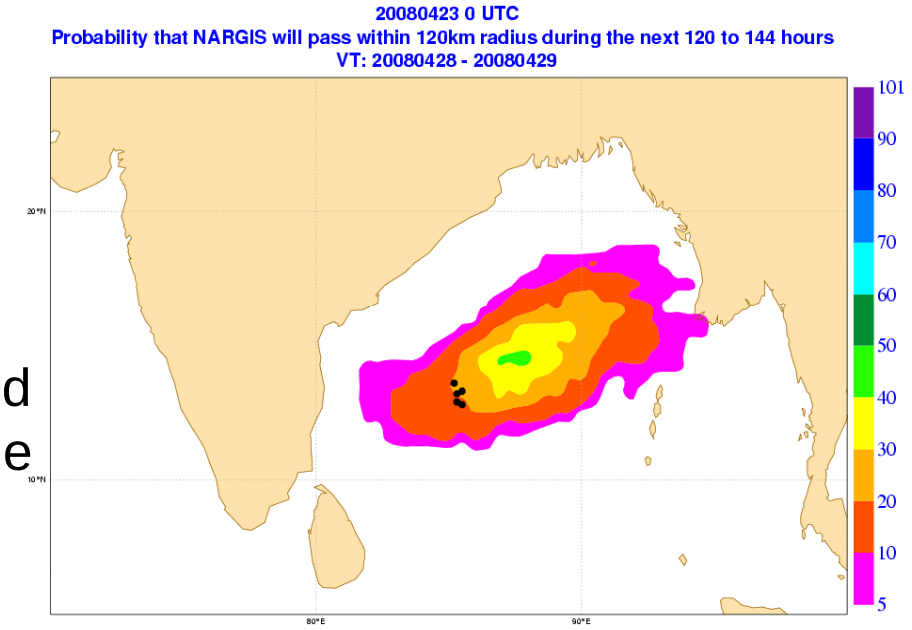
D+5 from 23 april 00z (28 april)



D+9 from 23 april 00z (2 may)



Black dots are the first official reported positions of cyclone (~april 28th)



A useful signal 9 days before the event

EPS probabilities from 23 april 00z at D+5 (28 april)

Where did the progress come from?

The development of the codes is highly mutualized

- **The IFS and Arpege codes have been developed jointly by ECMWF and Météo-France for the last 20 years**

- **The « ALADIN » consortium of Eastern and Southern Europe countries has joined the development 15 years ago**

- **The « HIRLAM » consortium of Western and Northern Europe countries has joined 5 years ago**

  - **About 30 european countries are now working together on this development effort, only few countries are still out**

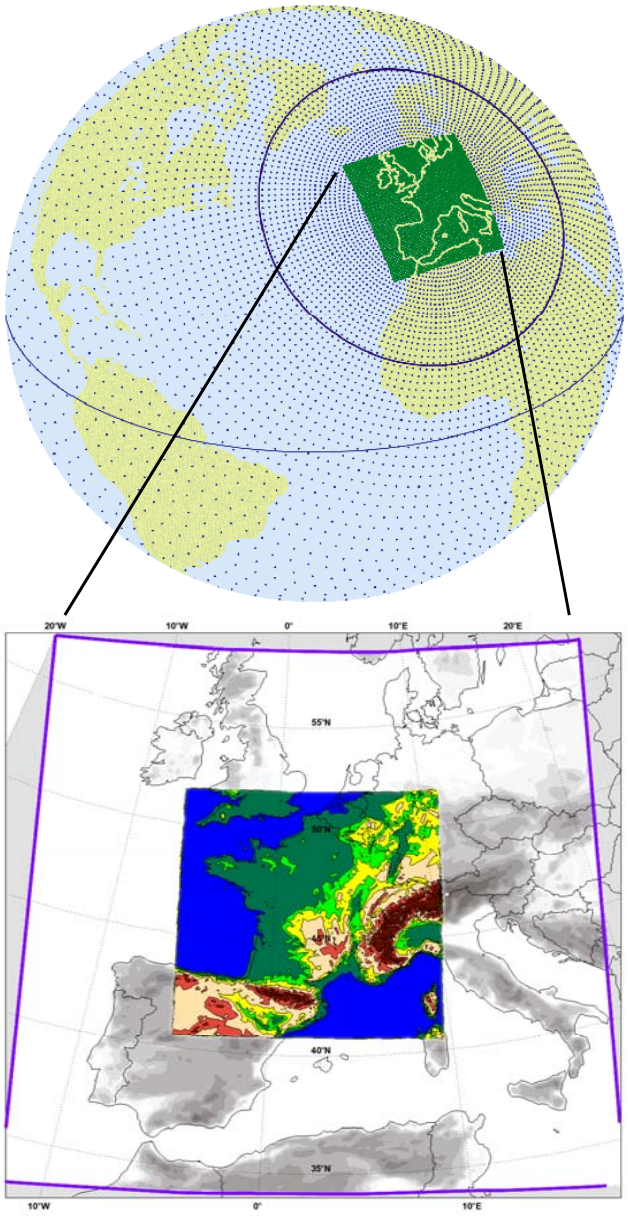
The exploitation is done in complementary configurations

- **ECMWF:**
  - **Global, uniform, high-resolution grid (25km) for deterministic medium-range forecasts (to 15 days ahead)**
  - **Global, uniform, moderate-resolution grid (50km) for probabilistic medium-range forecasts (to 15 days ahead) and monthly forecasts (80km, to 32 days ahead)**

- **Météo-France**
  - **Global, variable resolution grid (15km to 100km) for short-range forecasts (to 3 days ahead)**

- **Météo-France and other european countries**
  - **Limited-area domains at very high resolution (2.5 to 10km) for very short-range forecasts over Europe only (to 1 day ahead)**

The operational suite of forecast models at Météo-France

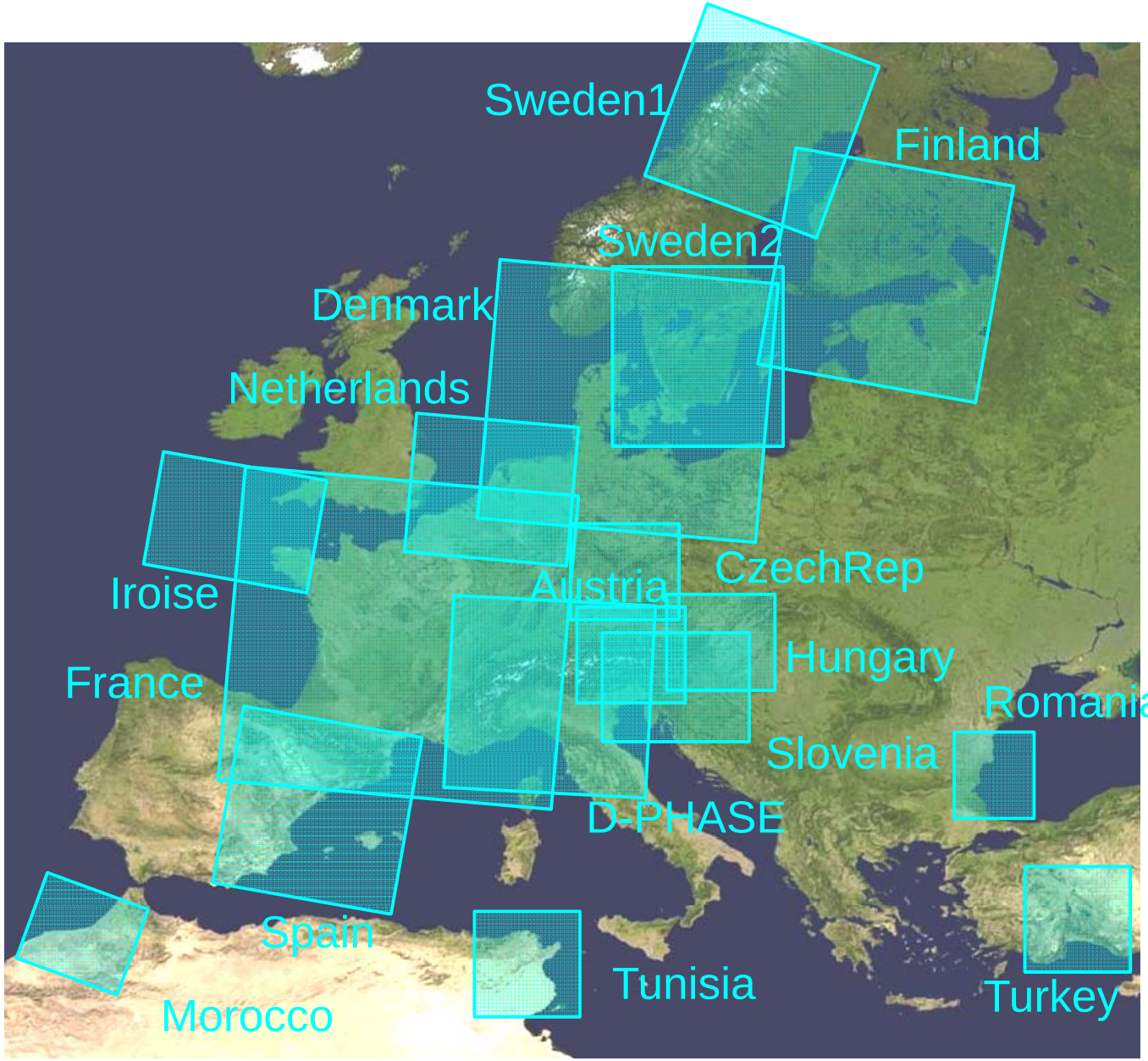


ARPEGE
Resolution 15km
Forecasts to 72 hours ahead

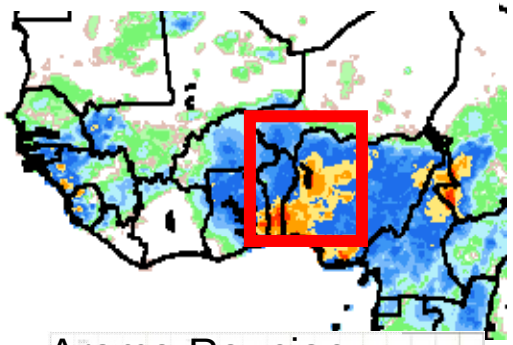
ALADIN (10km)
Forecasts to 48 hours ahead

AROME (2,5km)
Forecasts to 30 hours ahead

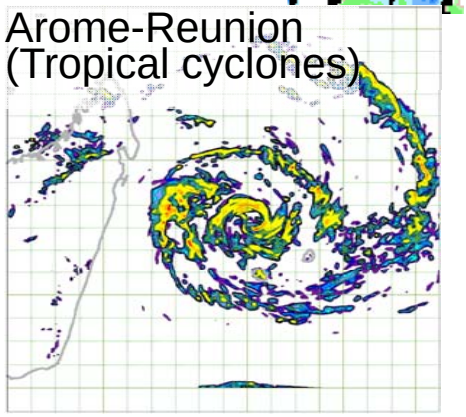
AROME domains under consideration in various countries



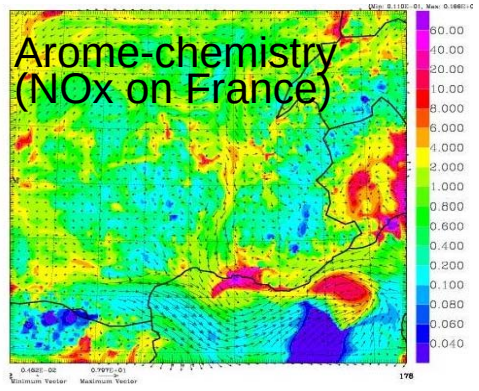
Arome-AMMA
(Western Africa)



Arome-Reunion
(Tropical cyclones)



Arome-chemistry
(NOx on France)



The numerical efficiency and accuracy of models have improved tremendously

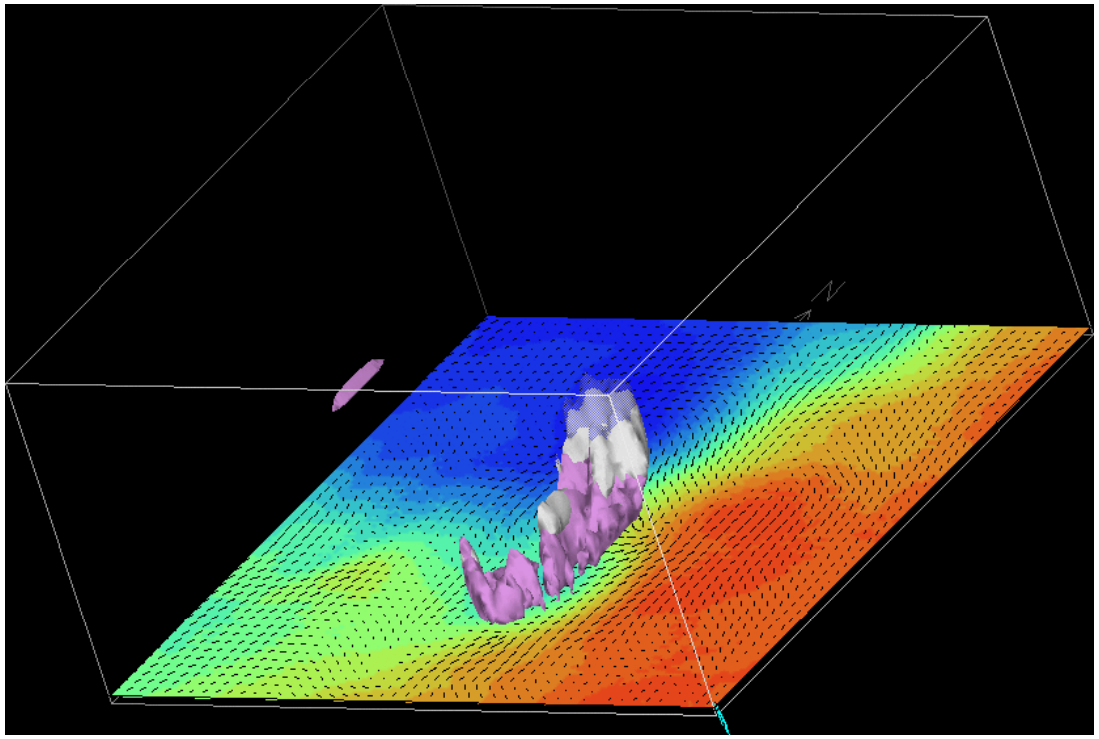
- **Reduced Gaussian grid and linear grid (spectral models)**
- **Implicit physics parametrization schemes**
- **Semi-Lagrangian advection**
- **Two-time level schemes (vs three-time level)**
- **Finite elements on the vertical**
- **Algorithmic optimization**

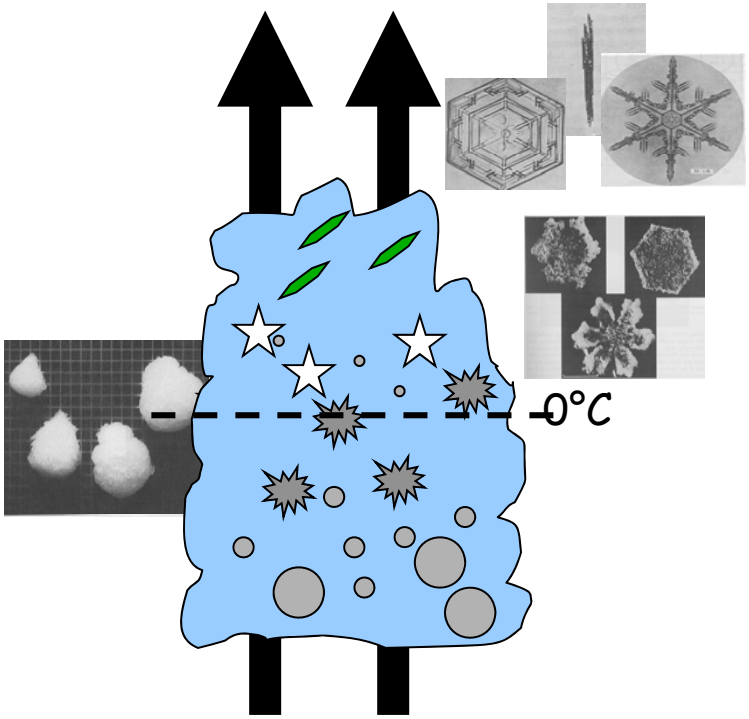
- **Overall gain: probably close to a factor 20**

Exemple of progress on the physics: better description of cloud microphysics and their impact on dynamics

Evaporation of rain drops creates massive cold pools of air that drive the dynamics of thunderstorms over flat terrain

(Simulation with AROME)





Increase of computing power and resolutions: Exemple of ECMWF

Year	Sustained TFlops	Headline global resolution (km)
1990	0.001	130
1995	0.01 (x10)	62
2000	0.3 (x30)	40
2005	2.5 (x8)	25
2010	20 (x8)	16
2015	160 (x8) ??	10 ??

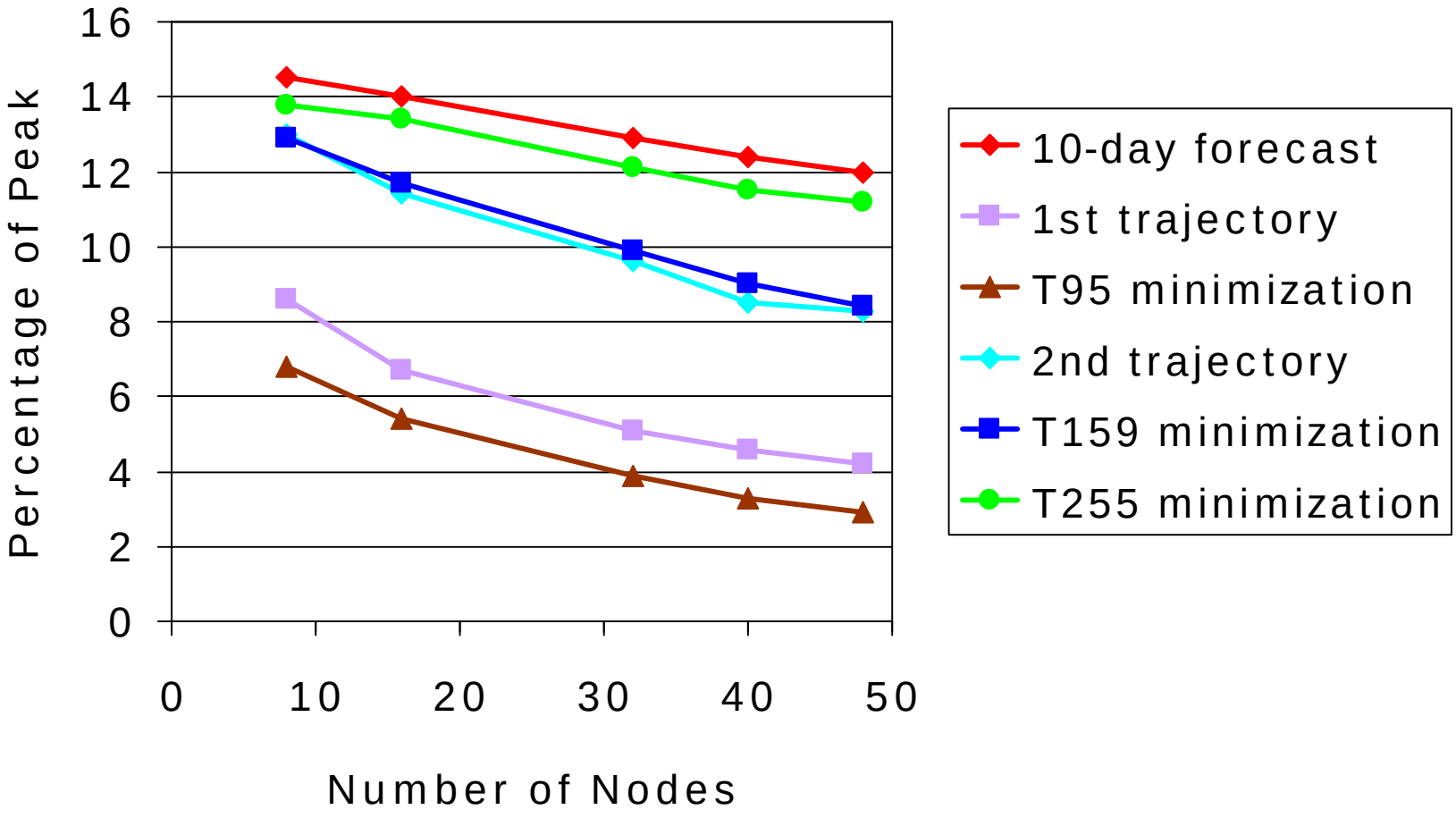
HPC money stream increased by 50% from 2009

Next money stream to be decided in 2011

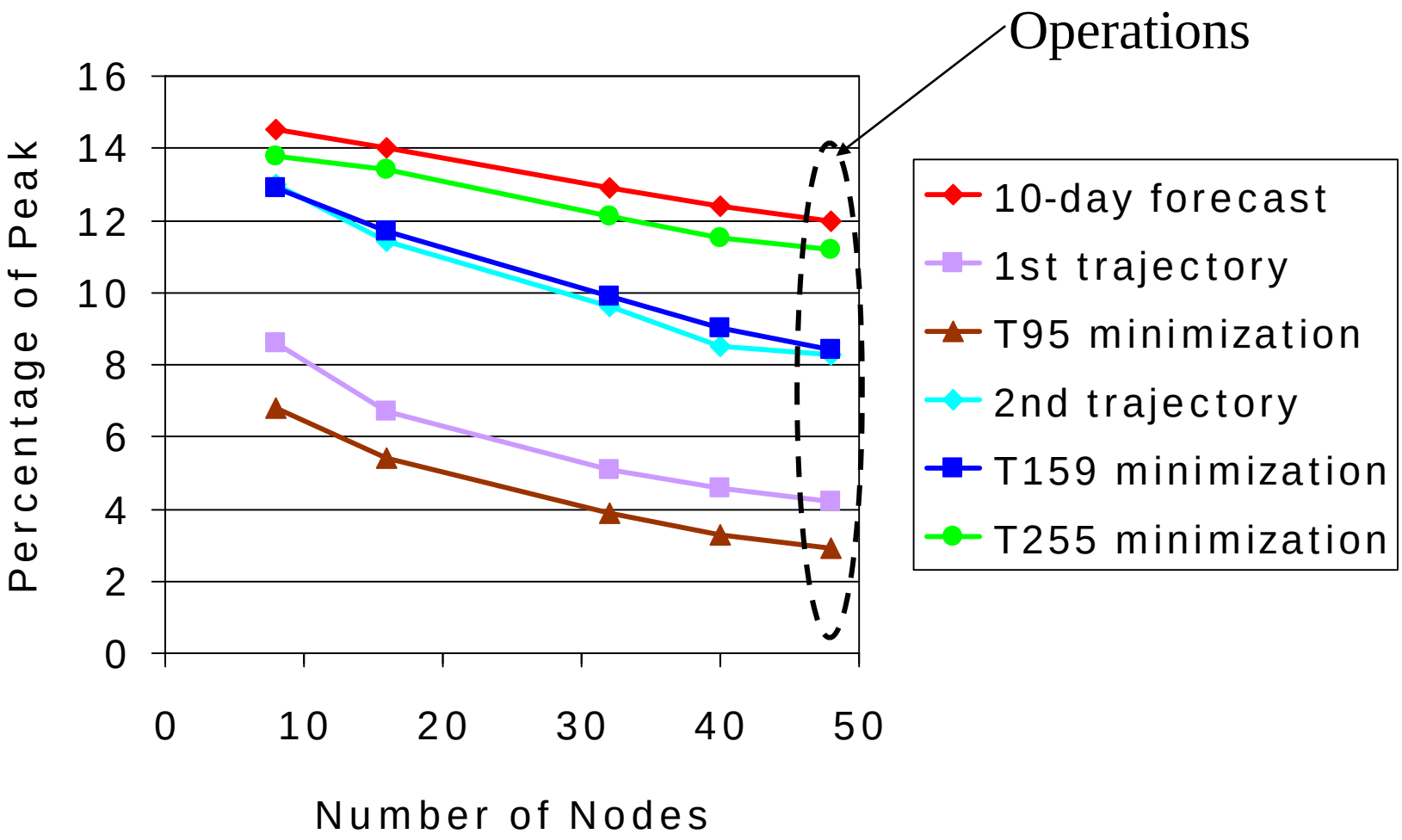
IFS parallelization: MPI & OpenMP

- **IFS-Arpege is maintained on both scalar and vector architectures**
- **Outer level: MPI tasks**
  - **For computations with distributed memory**
  - **Geographic partition for computations in physical space (physical aspects) – needs a large Halo**
  - **Transpositions to Fourier and spectral space for dynamics**
  - **Long messages and few synchronization points**
- **Inner level: OpenMP threads**
  - **For computations with shared memory within an MPI task**
- **The number of MPI tasks and OpenMP threads is optimized for each model configuration**
  - **Typically 192 tasks x 8 threads for operational jobs**

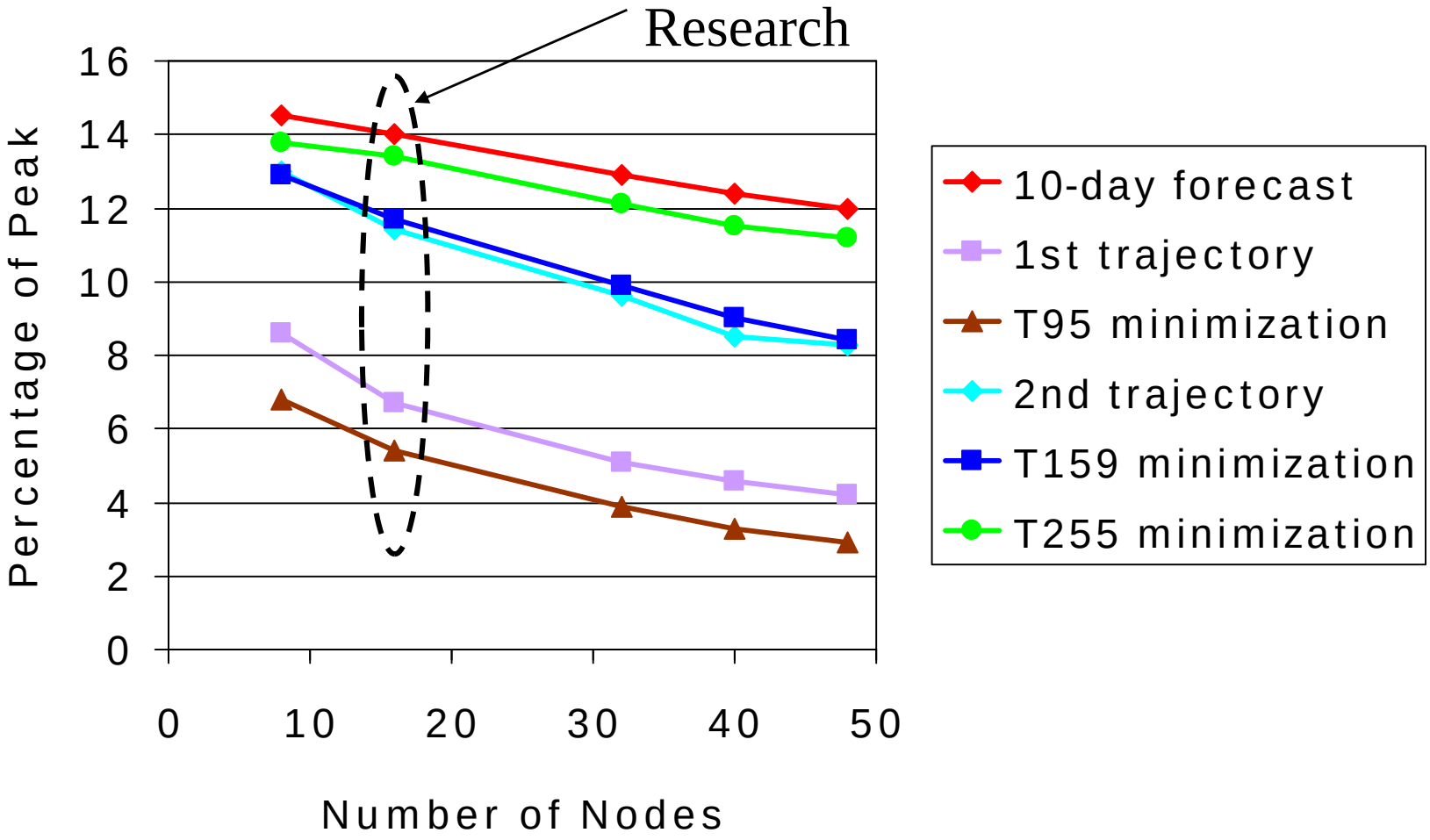
Percentage of peak power used by various applications



Percentage of peak power used by various applications

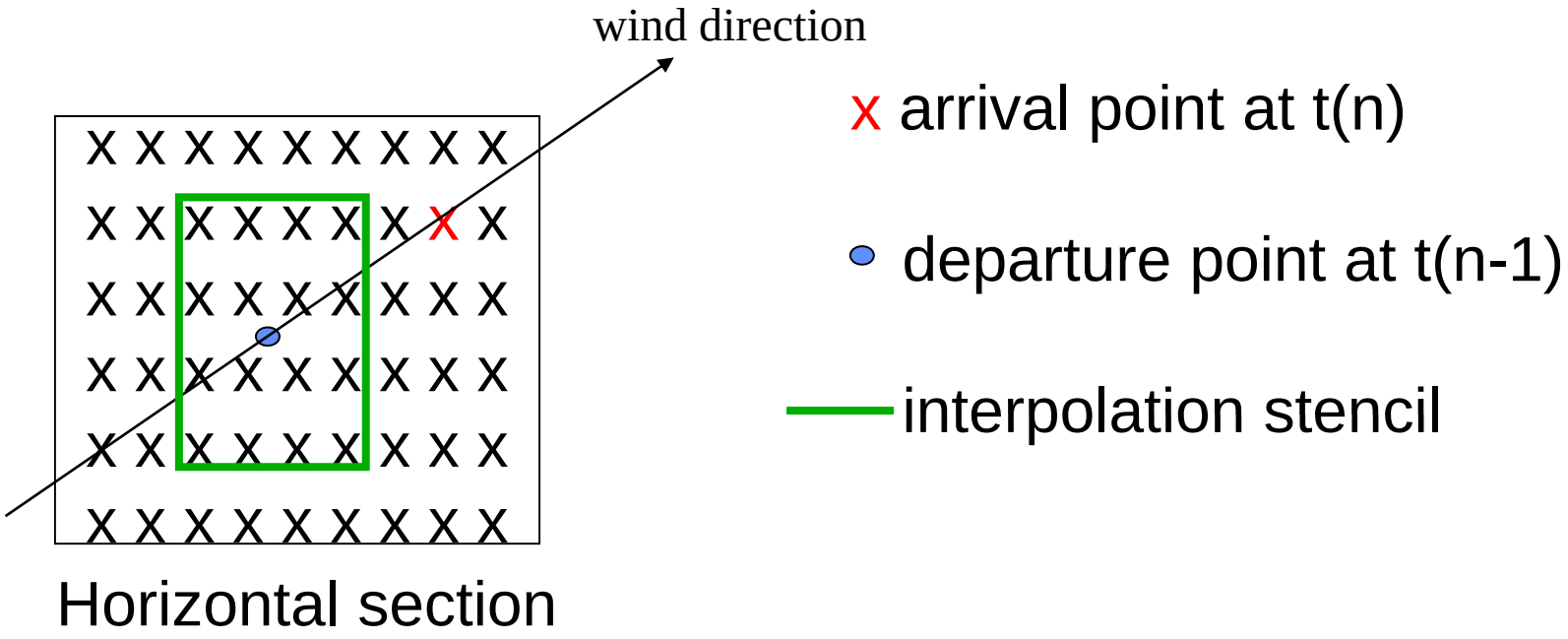


Percentage of peak power used by various applications



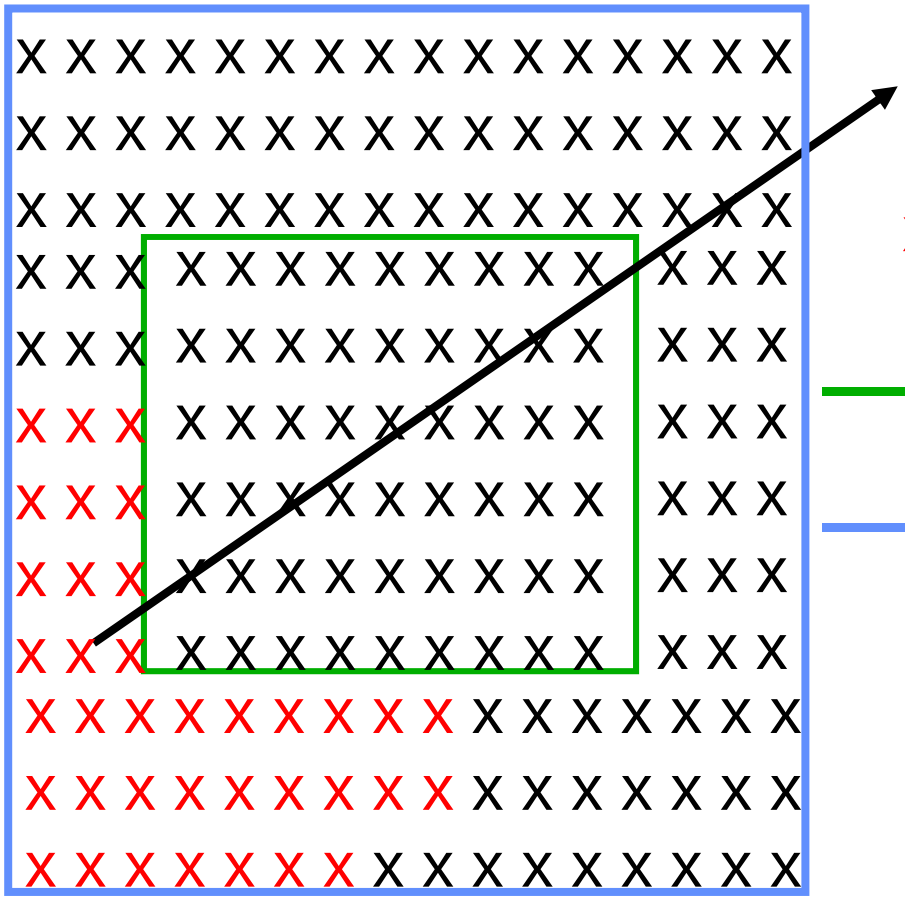
Examples of recent optimizations

IFS – Semi-Lagrangian Advection



Full interpolation in 3-D is 32 point

IFS – Semi-Lagrangian 'Halo on Demand' (Deborah Salmond)

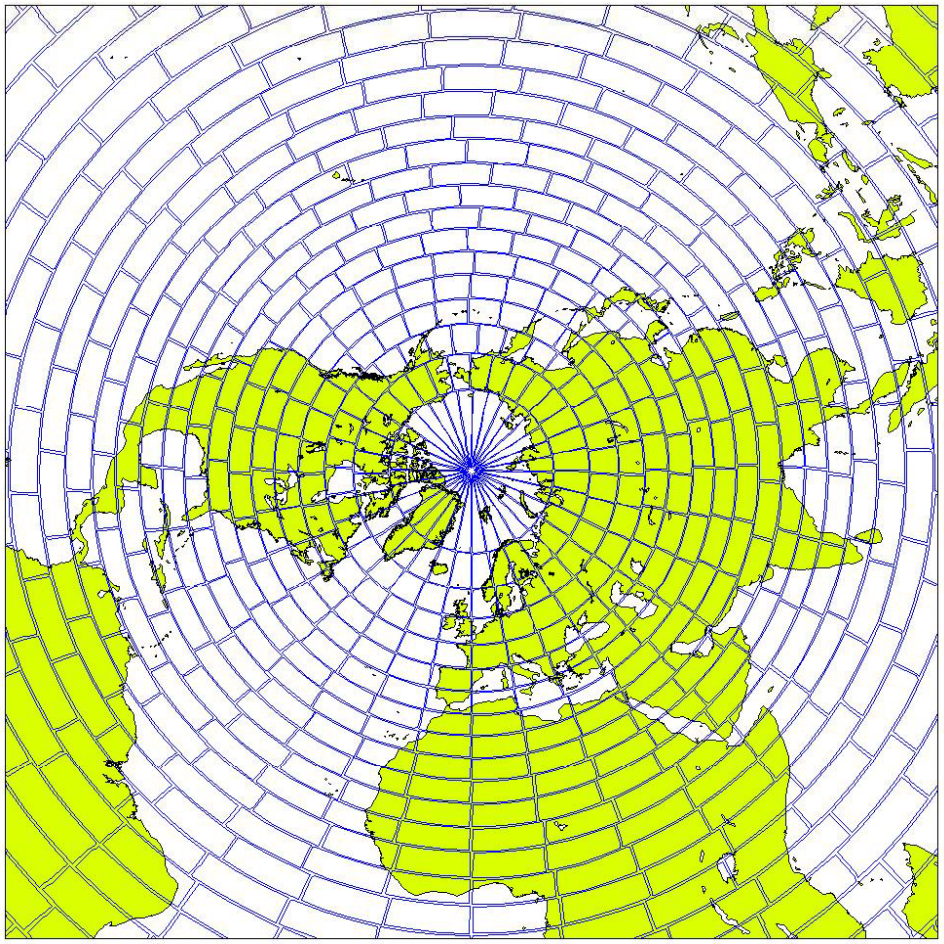


x points needed in halo

points located on task n

halo

Improvement of domain decomposition for MPI tasks (George Mozdzinsky)



2D partitioning results in non-optimal Semi-Lagrangian comms requirement at poles and equator!

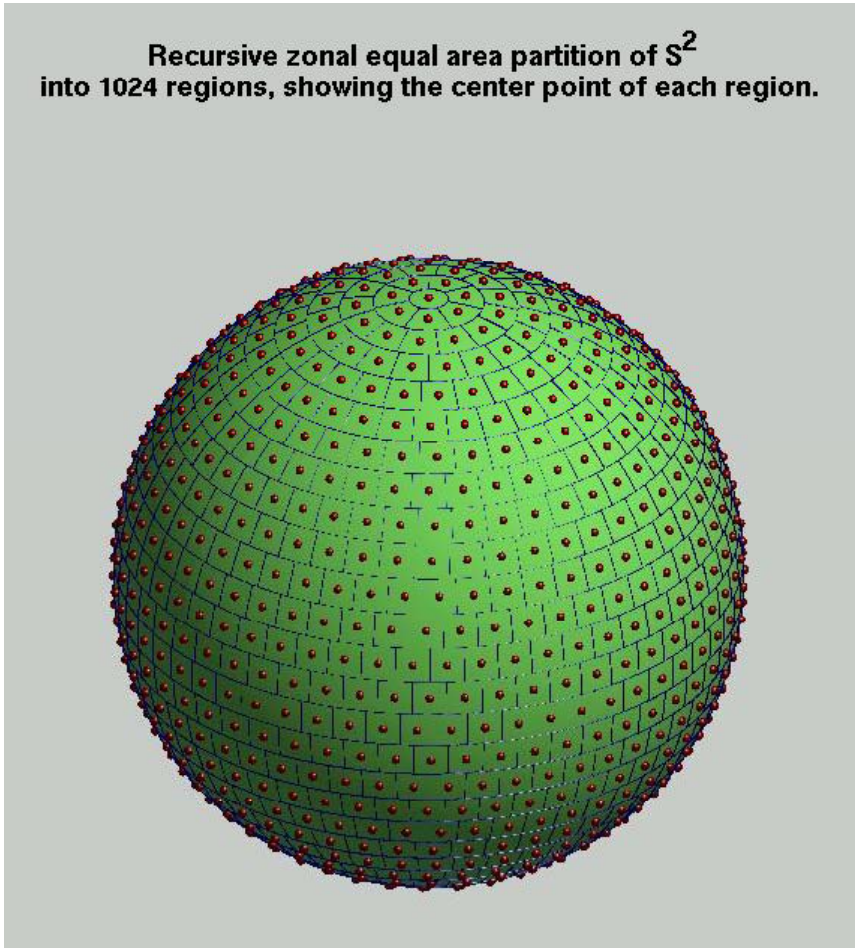
Square shaped partitions are better than rectangular shaped partitions for dealing with halos

An elegant solution: the eq_regions algorithm

Developed by Paul Leopardi et al. , School of Mathematics, Univ. of New South Wales, Australia.

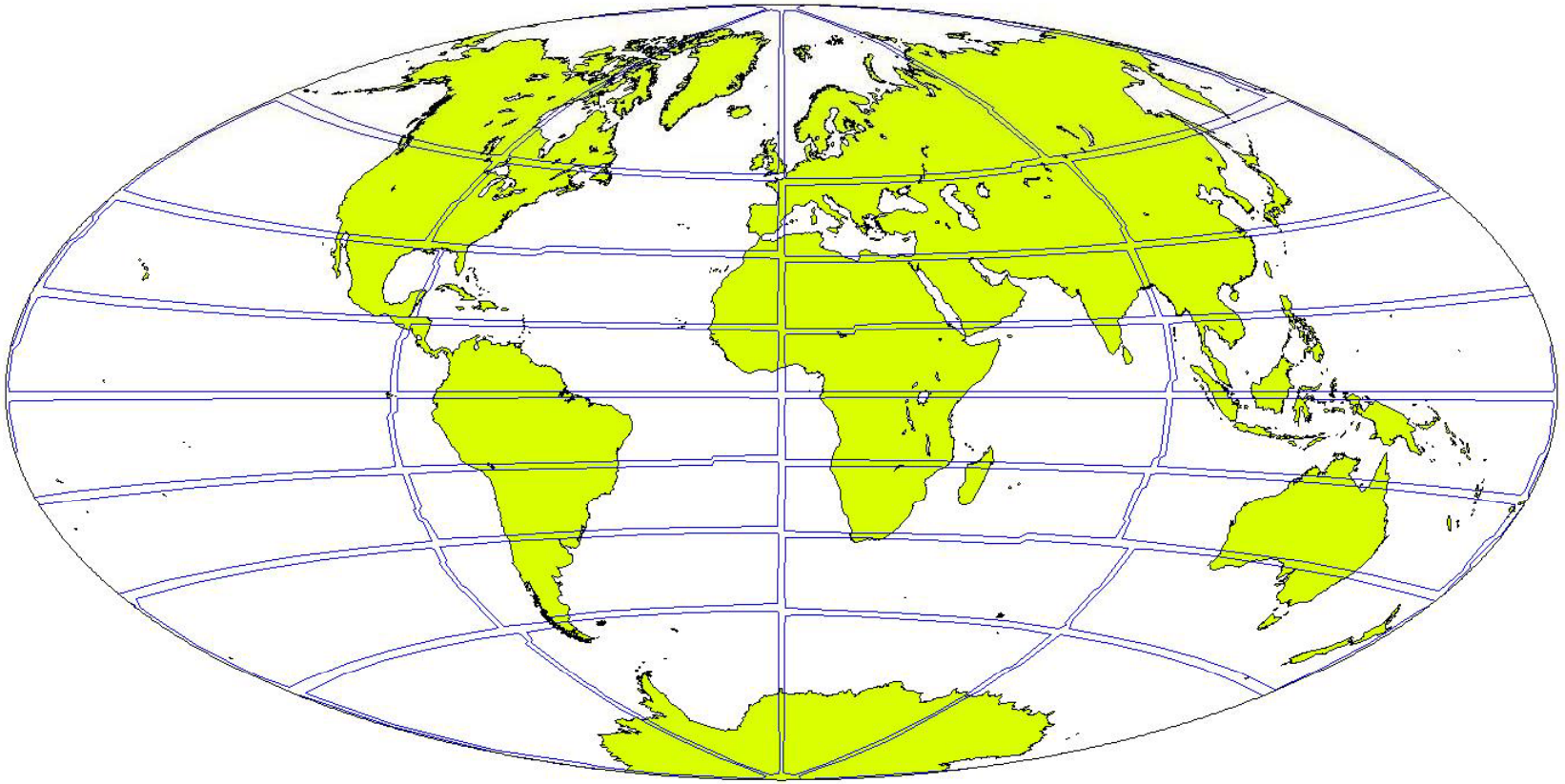
Paper:

http://www.maths.unsw.edu.au/applied/files/2005/amr05_18.pdf

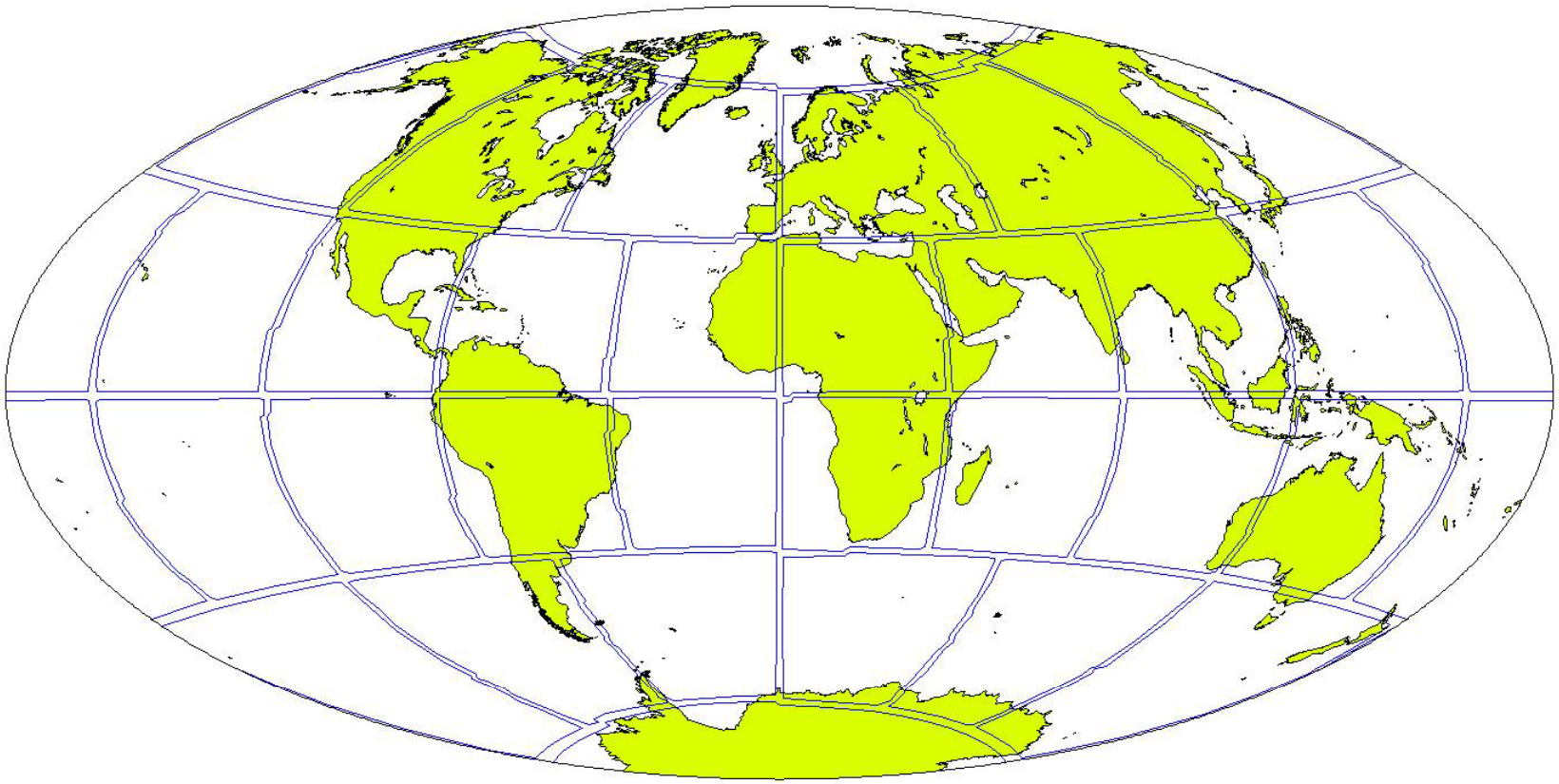


2D partitioning T159 32 tasks (NS=8 x EW=4)

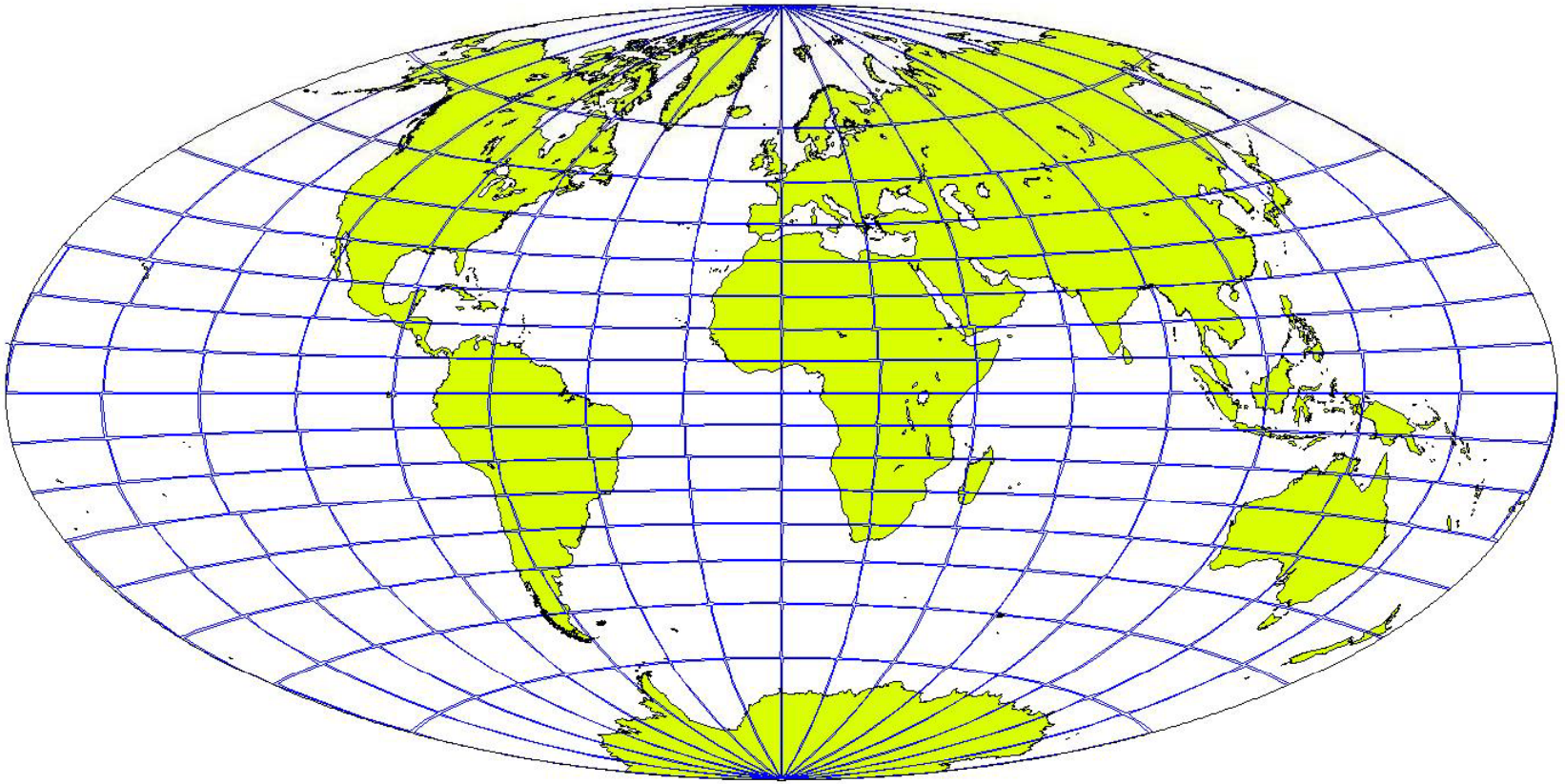
Aitoff projection



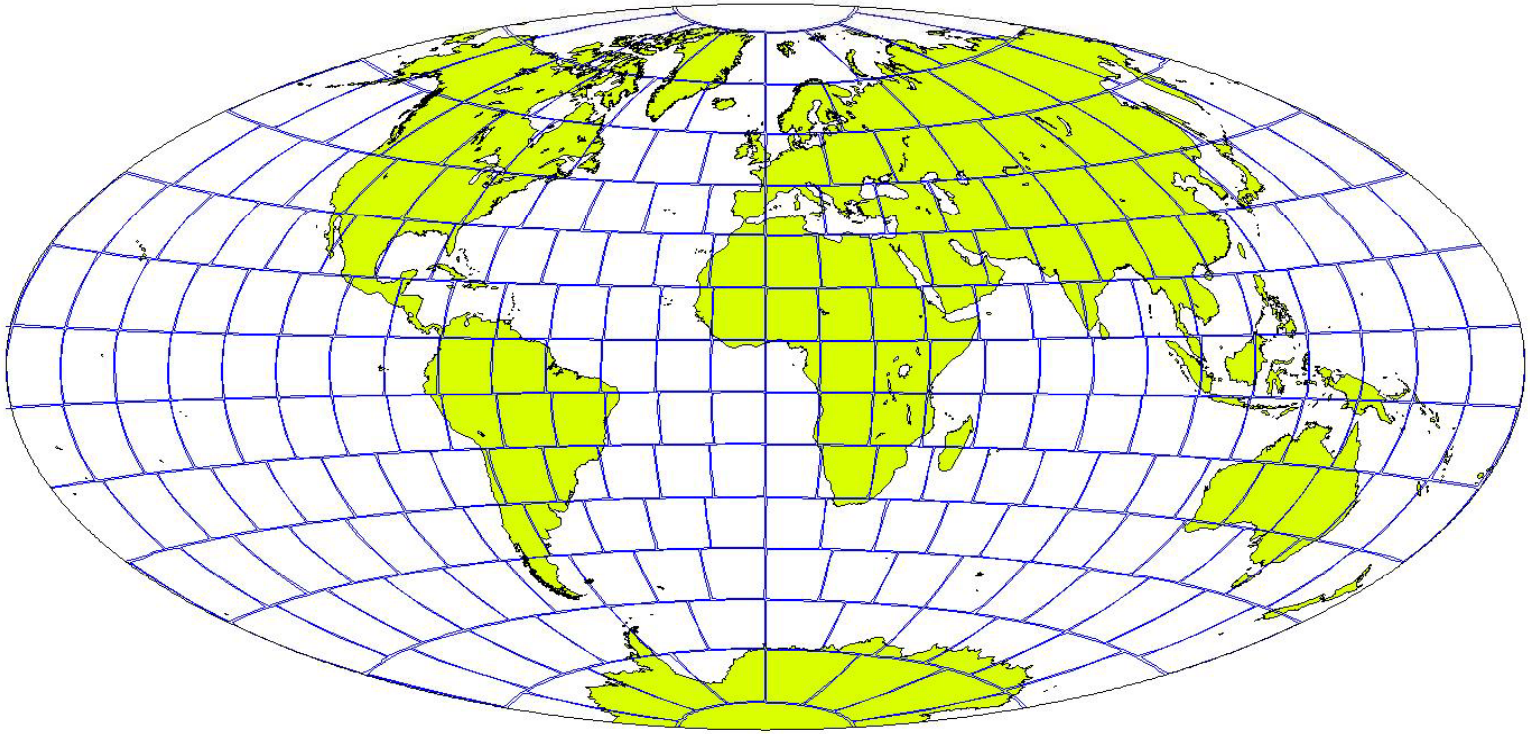
eq_regions partitioning T159 32 tasks



2D partitioning T799 256 tasks (NS=16 x EW=16)



eq_regions partitioning 256 tasks (data assimilation)
Improvement of 2.7% in efficiency



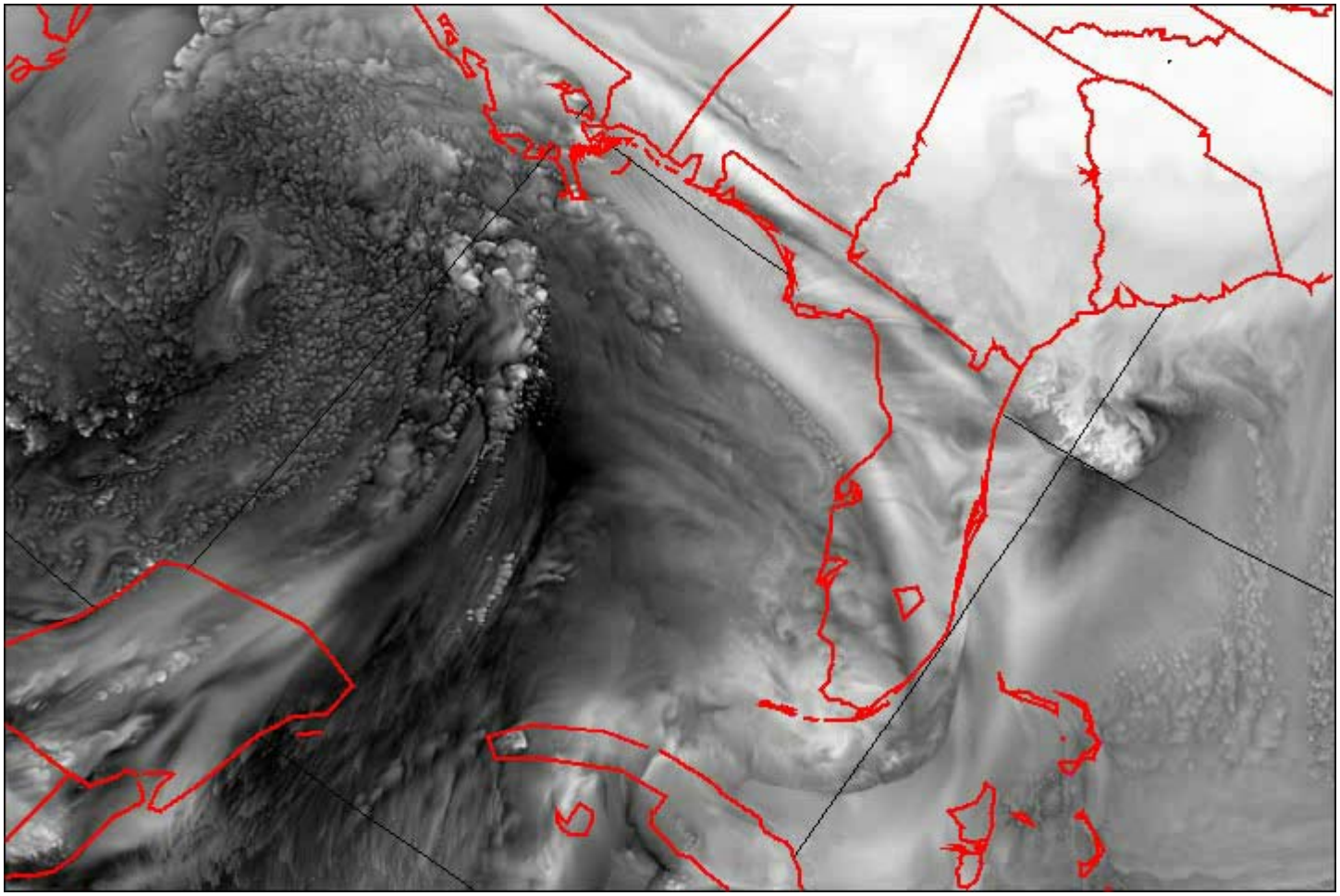
Current and planned global resolutions (km) at ECMWF and other centres (Source: WMO)

	2008	2009	2010	2011
ECMWF	**25**	**16**	**16**	**16**
UK	**40**	**25**	**25**	**20**
France	**16-90**	**16-90**	**10-60**	**10-60**
Germany	**40**	**20**	**20**	**15**
USA	**50**	**50**	**25**	**25**
Canada	**35**	**35**	**25**	**25**
Japan	**20**	**20**	**20**	**20**

Current and planned regional resolutions (km) at national centres (Source: WMO)

Country	2008	2009	2010	2011
UK	4	1.5	1.5	1.5
France	2.5	2.5	2.5	2.5
Germany	2.8	2.8	2.8	2.8
USA	4	4	4	3
Canada	15	10	10	tbd
Japan	5	5	5	5

1km resolution on very large domains (target for ????)



Simulation by Gilbert Brunet, Canadian Met Service, on the Earth Simulator

When shall we need a major reorganization of the IFS-Arpege code?

- **Currently the number of MPI tasks is 200 – 1000 and corresponds exactly to the number of subdomains**
- **The number of cores attributed to each subdomain is much smaller than the number of grid points**
- **In the future it is likely that both the number of subdomains and the number of cores per subdomain will increase dramatically**
- **Increase in the number of subdomains will create a growing overhead of communications, especially as the size of the halo will not decrease**
- **Will the number of cores per subdomain approach or exceed the number of grid points? This would require a complete reorganization of the code**

The Climate Problem

- **World Modelling Summit for Climate Prediction was organized by WCRP and hosted by ECMWF 6-9 May 2008**
- **See all presentations at http://www.ecmwf.int/newsevents/meetings/workshops/2008/ModellingSummit/presentations**

Key ideas discussed at the Modelling Summit

- **Climate models would really benefit from increased resolution**
  - **Systematic errors improve dramatically when resolution increases (current range 300km - > 100km)**
  - **First global 3.5km resolution climate run achieved by Masaki Satoh (JAMSTEC) on the Earth Simulator shows very encouraging results in capturing the variability tropical clouds**

- **Predicting climate also requires other aspects**
  - **More complexity (ocean, land surface, biogeochemistry, etc…)**
  - **Long simulations (1000 years?)**
  - **Ensembles**

Key ideas discussed at the world modelling summit

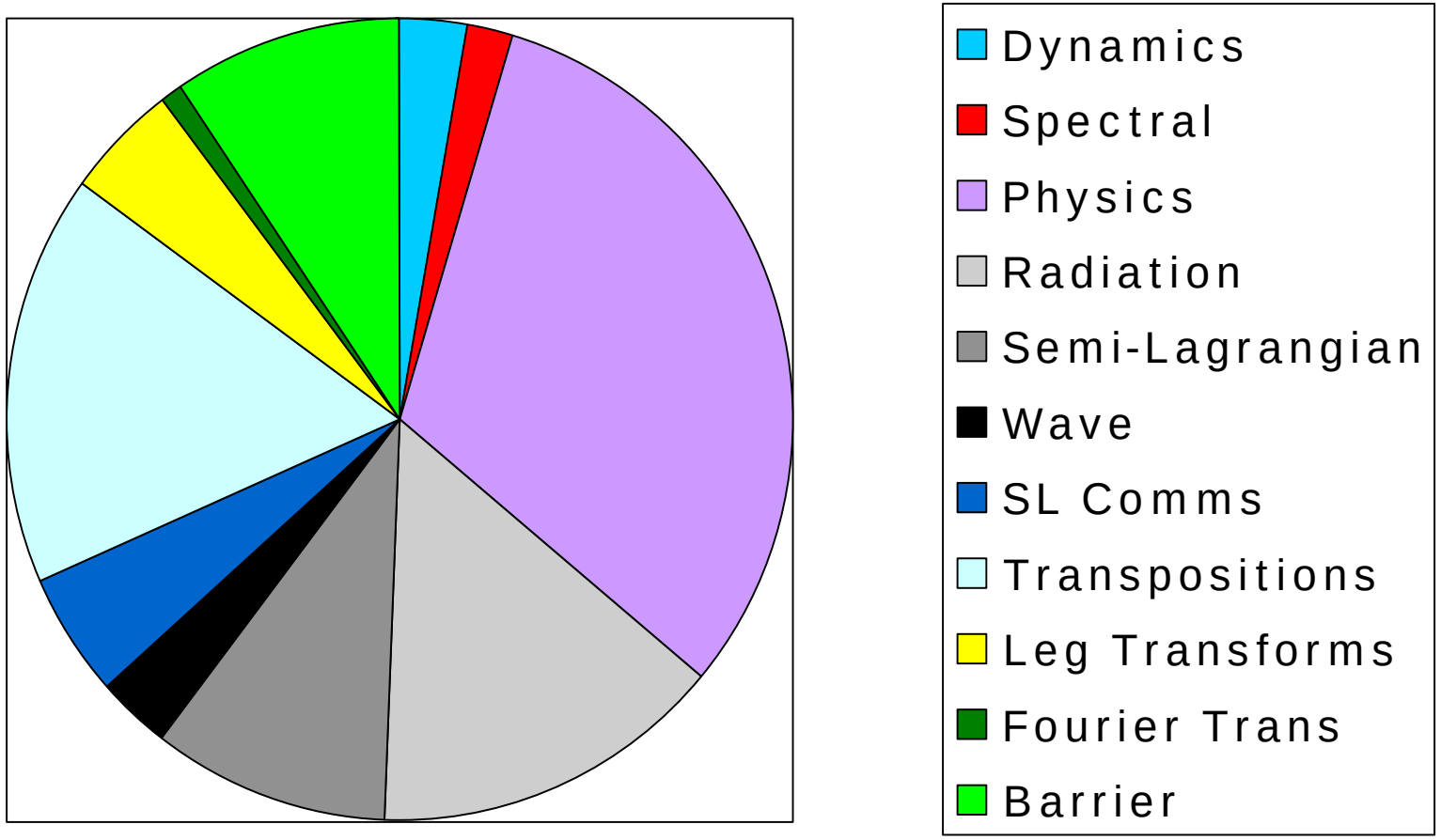
- **Resolution of climate models may catch up and even surpass the resolution of weather forecast models . It makes sense to use the same models**
  - **Avoid duplication of cost of development**
  - **Benefit from validation in "weather forecast mode" to eliminate dubious formulations and reduce uncertainty in climate prediction**

- **The Summit "Statement" advocated**
  - **Developing a synergy between Weather and Climate**
  - **Creating a new "World Climate Research Facility", featuring one or several machines, several orders of magnitudes larger than the current largest computers**
  - **Starting new efforts in adaptation of climate models to future computer architectures**

Thank you

T799 L91 10-day Forecast – 32 Nodes

Where the time is spent



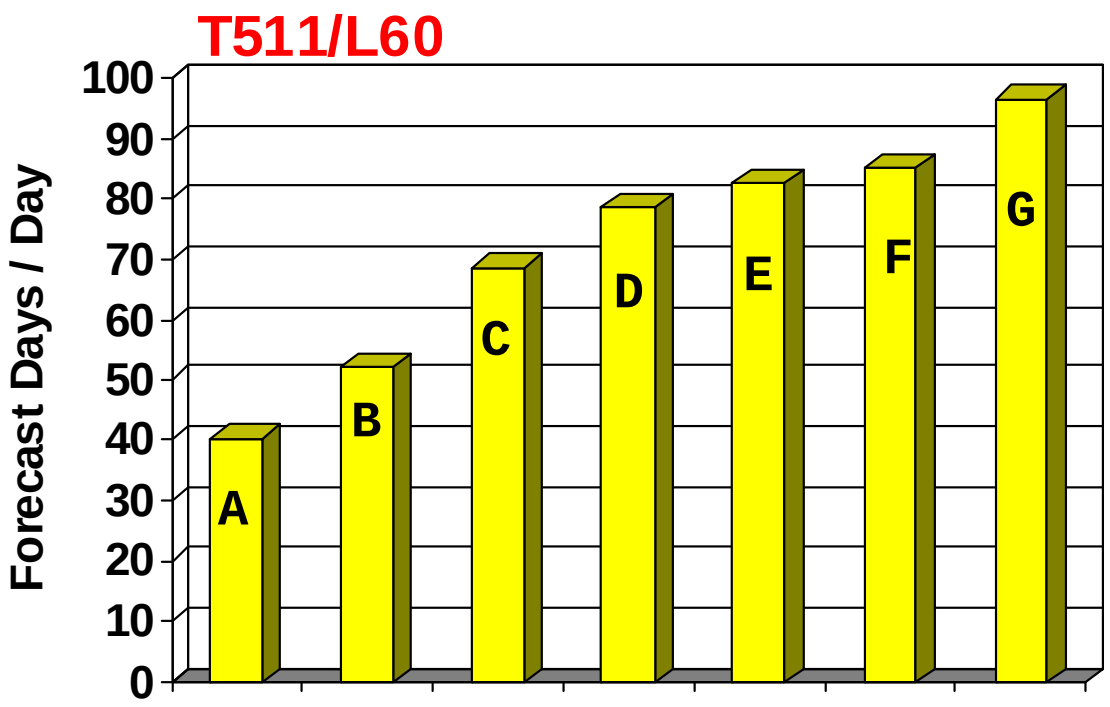
Comparison of the new and old HPCF

	New Phase1 HPCF	Current HPCF
No. of clusters	2 compute clusters 2 I/O storage clusters 1 small test cluster	2 compute clusters 1 MC-GPFS “controlling” cluster 1 small test cluster
Performance	~ 20 TFLOPS (sustained)	~ 4 TFLOPS (sustained)
Each compute cluster		
Operating System	AIX 5.3 (probably, but AIX 6.1 is possible)	AIX 5.3
Compute nodes	248 x 32-core POWER6 (SMT)	155 x 16-core POWER5 (SMT)
Compute processors	~8000	~2500
Network nodes	9 x 32-core POWER6 (connected to the LAN and I/O storage fabric)	2 x 16-core POWER5 (connected to LAN)
I/O (VSD) nodes	None	8 x 16-core POWER5 (connected to the fibrechannel SAN)
Interconnect	8-plane IB4x-DDR	2-plane pSeries HPS (Federation)
I/O subsystem		
Paradigm	Independent I/O storage clusters	Fibrechannel SAN
Disk types	Directly attached RAID6 storage	FAStT900 RAID5 subsystems
Disk space	1.8 Petabytes in total (1.3 PB initially)	100 Terabytes in total
Each compute server (node)		
Memory	64 Gigabytes (8 with 256 GB)	32 Gigabytes (4 with 128 GB)
Dual-core chips	16	8
Processors (cores)	32	16
Each processor (core)		
Lithography	65nm	90nm
No. of transistors	790 million	276 million
Clock frequency	4.7 GHz	1.9 GHz
Peak performance	18.8 GFLOPS (~290 TFLOPS total HPCF)	7.6 GFLOPS (~37 TFLOPS total HPCF)

Computing requirements for some global configurations of IFS-Arpege

Resolution	T1279 L91 Next Operational Resolution	T799 L91 Current Operational Resolution	T399 L62 Ensemble Prediction System
Grid spacing	16km	25km	50km
Number of grid-points (horizontal)	2,140,704	843,490	213,988
Time-step	450 secs	720 secs	1800 secs
Floating-point ops for 10-day forecast	$7.207*10^{15}$	$1.615*10^{15}$	$0.1013*10^{15}$ → EPS * 50

Impacts of various Optimisations on IFS Performance on IBM (some years ago)



A: Base (128 cores)

B: Size of vector loops

C: Semi Lag "on demand"

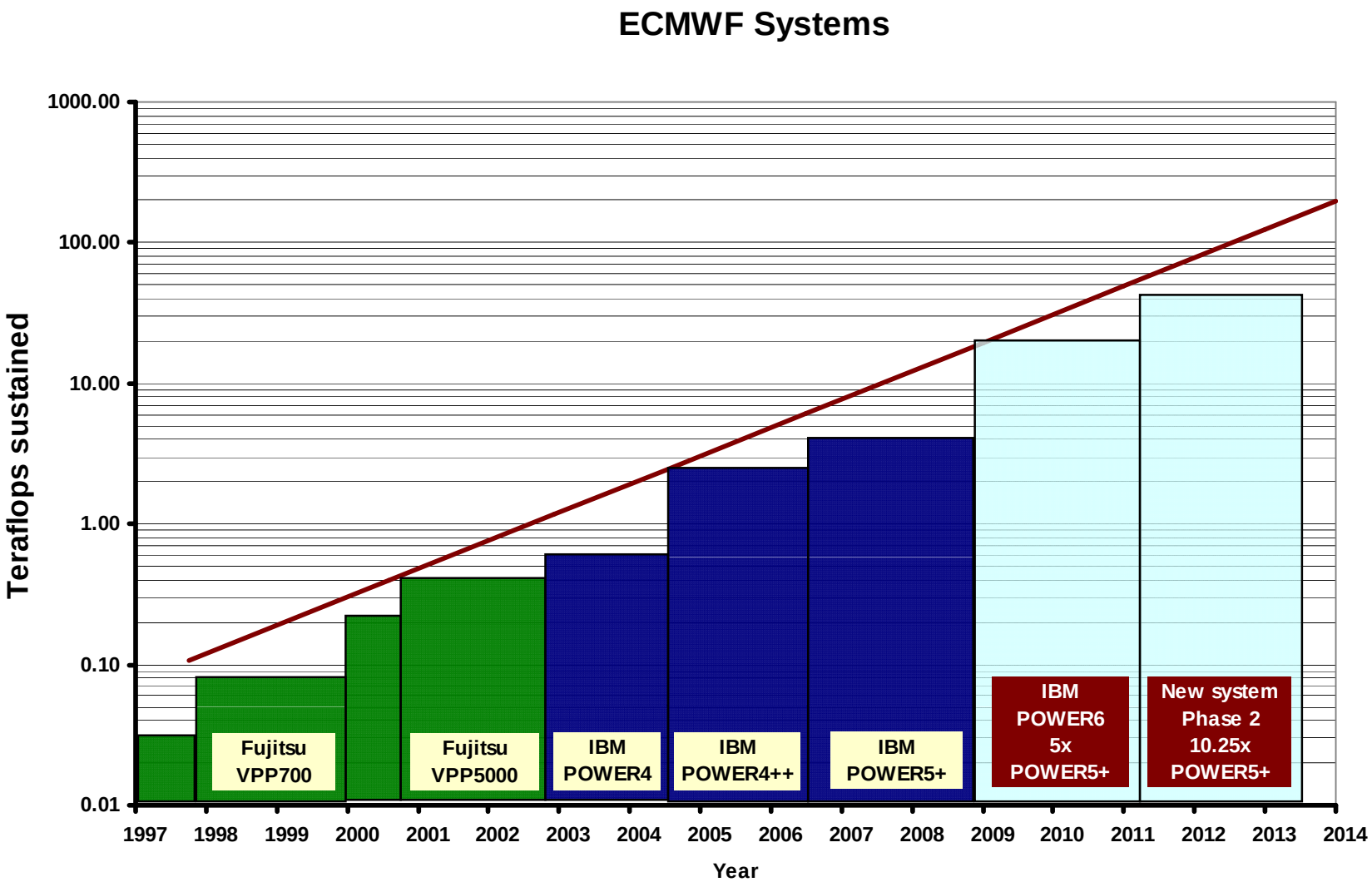
D: Domain decomposition

E: Comms not overlapped

F: IBM Vector Functions

G: OpenMP 1-> 2 Threads

A look into the future



The number and quality of satellite data used to constrain the initial state has increased dramatically