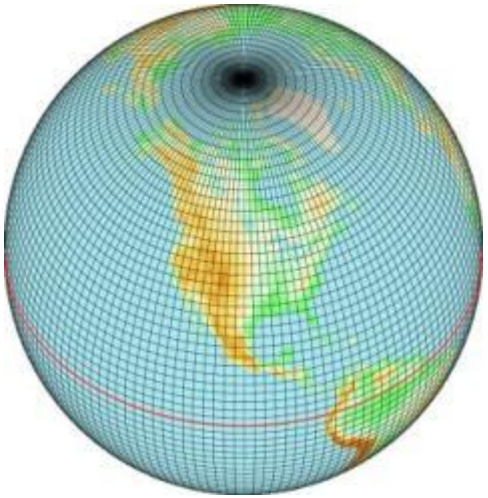
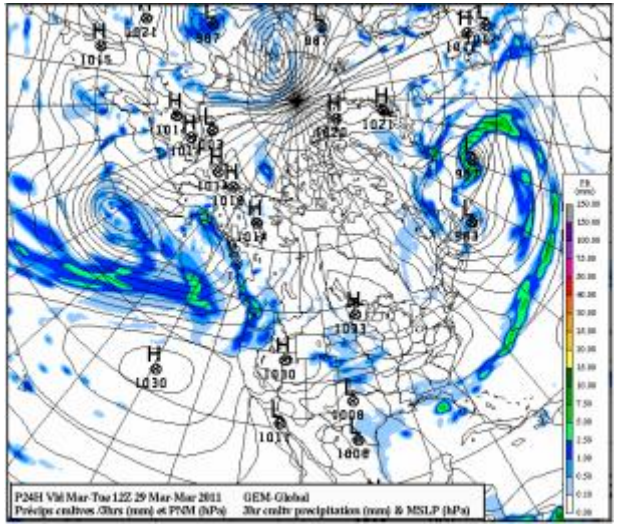
Environment Canada Environnement Canada

Canada

15th ECMWF Workshop on HPC in Meteorology
1-5 October 2012, Reading, UK

High Performance Computing enabling NWP at the Canadian Meteorological Centre





B. Denis, **Y. Chartier**, National Prediction Development Division
D. Bender, National Prediction Operations Division
L. Chardon, Numerical Prediction Research Division
A. St-Denis, HPC support

Outline

- Historical HPC evolution and forecast quality at CMC
- NWP systems running on EC supercomputer
- Next 10 years CMC NWP evolution
- New operational EC supercomputer IBM Power 7
- P5 to P7 model code migration
- Preparing the post-P7 era

Canadian Meteorological Centre

Meteorological Research Division:

Data Assimilation, Modeling, Cloud Physics

CMC Development Division:

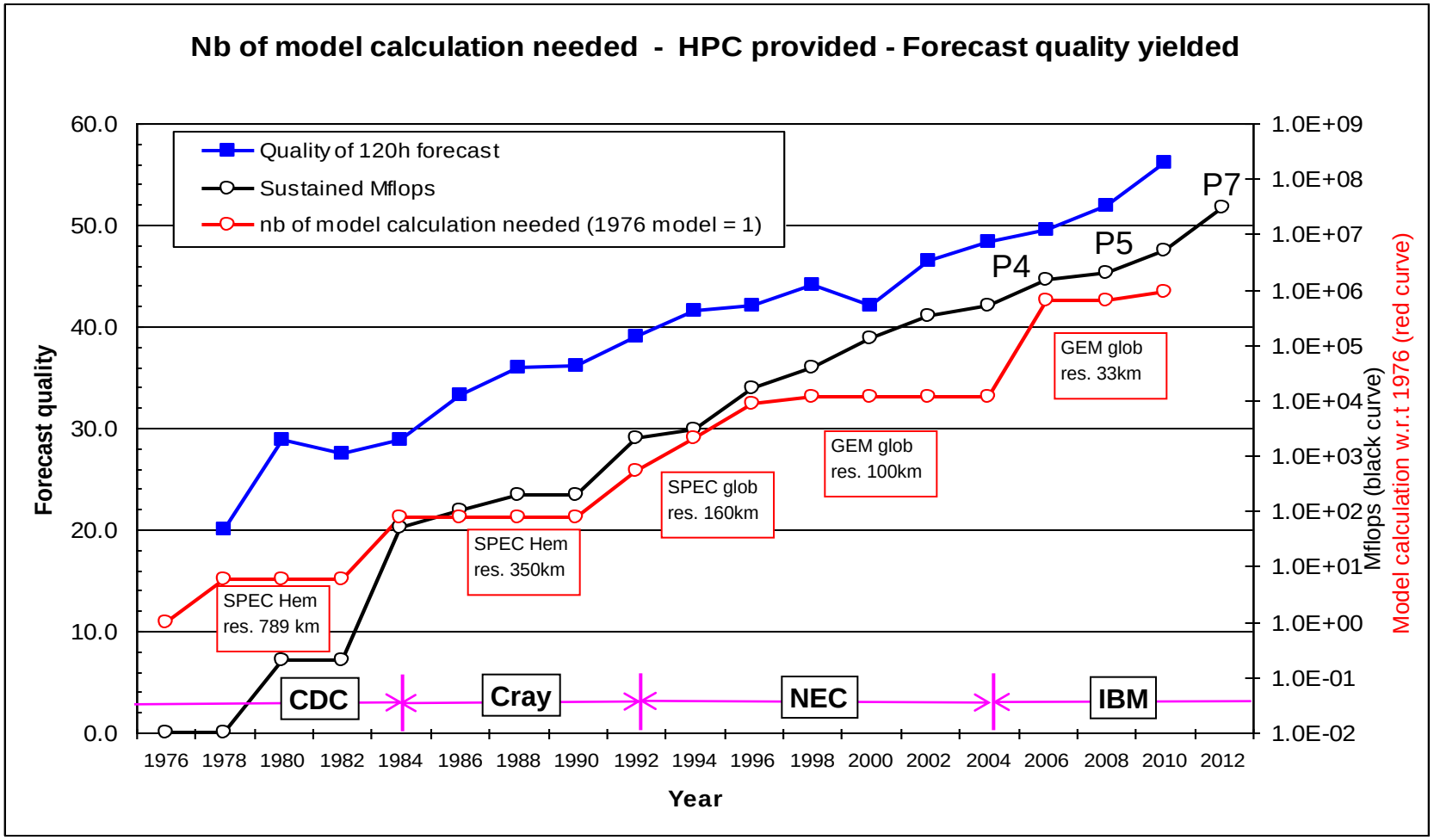
Data Assimilation, Numerical Weather Prediction, Weather Elements, Scientific Applications

IT Infrastructure:
Supercomputer, National Telecommunications, Network, User support

CMC Operations Division: Analysis &Prognosis, Env. Emergency Response, Air Quality, Implementation and Operational Services

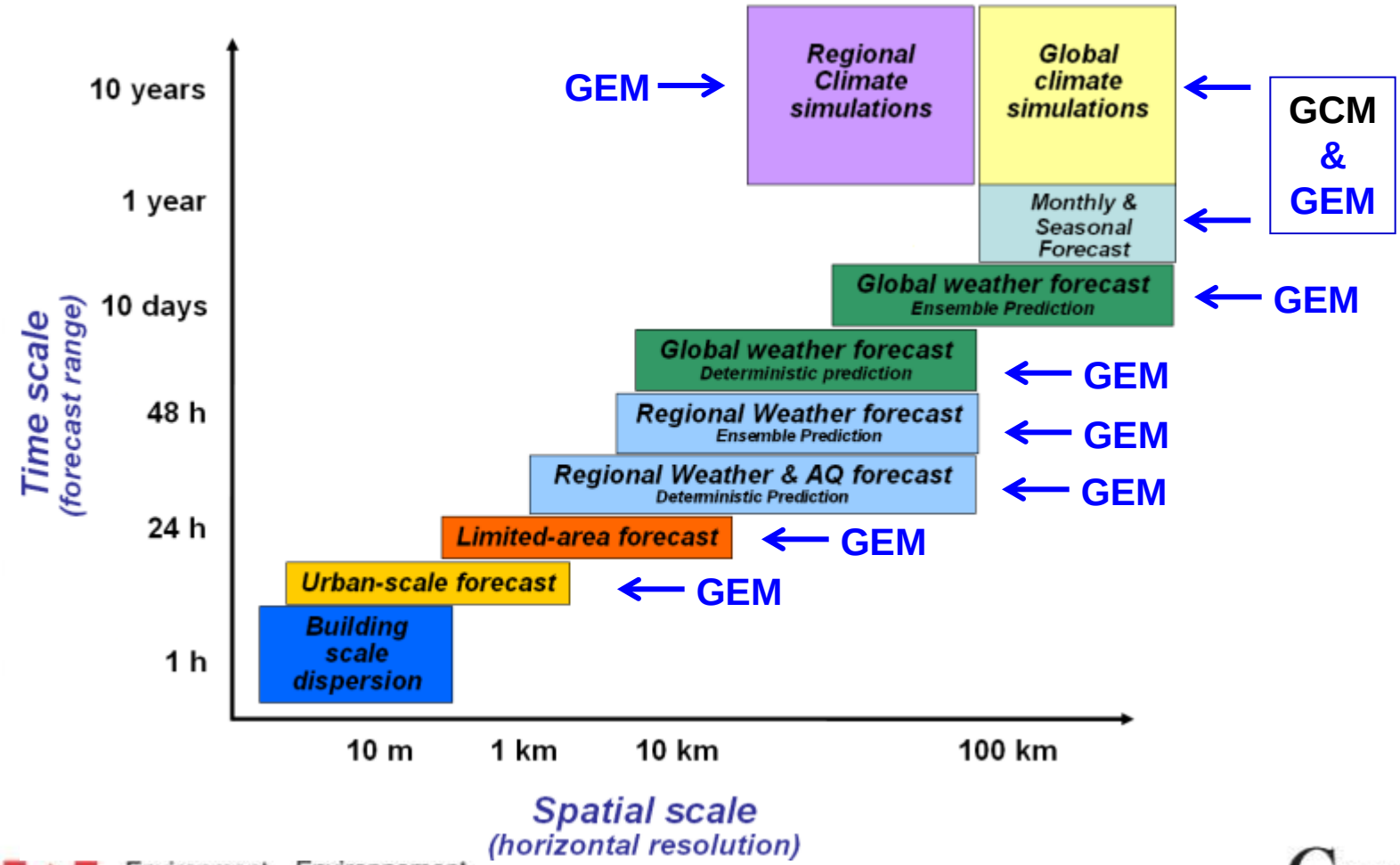
Canada

Historical HPC evolution and forecast quality at CMC



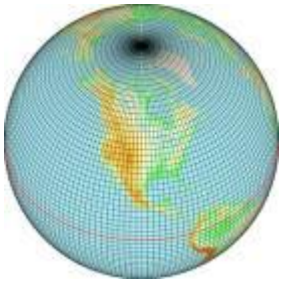
Modeling systems running on EC supercomputer

- GEM is the core model for many EC's applications

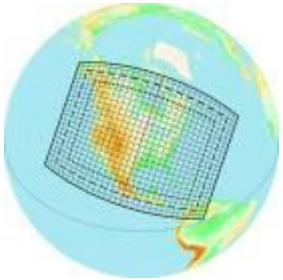


CMC-MRD unified multi-scale model : GEM

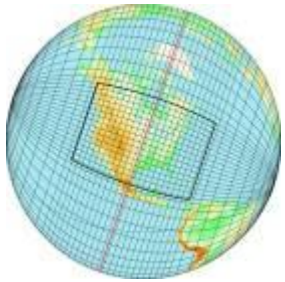
GEM= Global Environmental Multi-scale



- Global constant resolution (regular lat-lon grid) (GEM-Global)
  - Medium-range deterministic forecasts (33 km, soon 25 km)
  - Seasonal forecasts / Climate simulations (100 - 200 km)
  - Medium-range EPS (66 km)



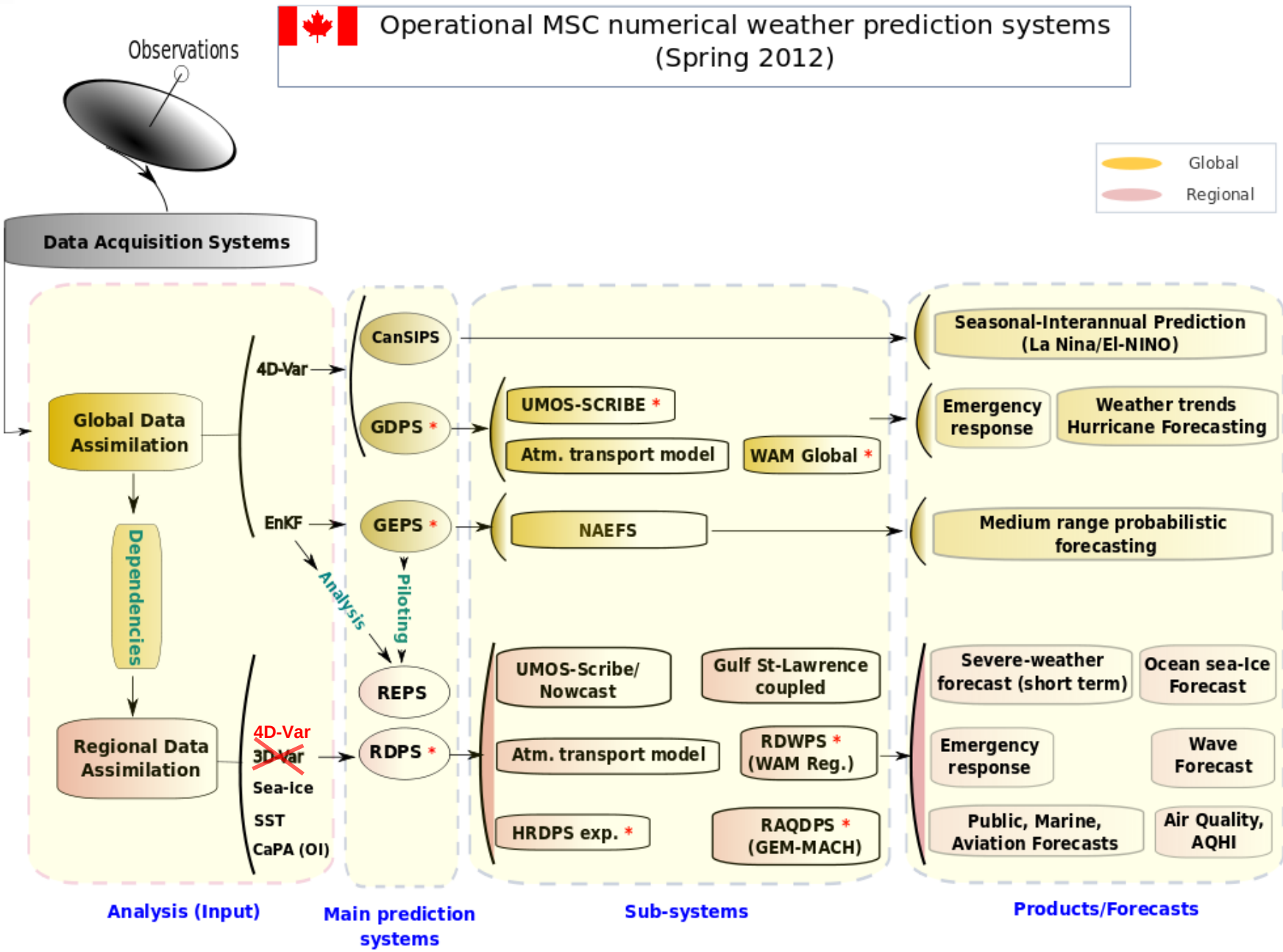
- Limited-area (LAM) constant resolution lat-lon grid
  - Short-range high-res. deterministic forecasts (GEM regional 15 km, soon 10 km and LAM 2.5 km)
  - Short-range EPS (33 km, soon 15 km)
  - Urban emergency response (250 m)
  - Regional climate simulations (15-55 km)



Global variable resolution (stretched grid)

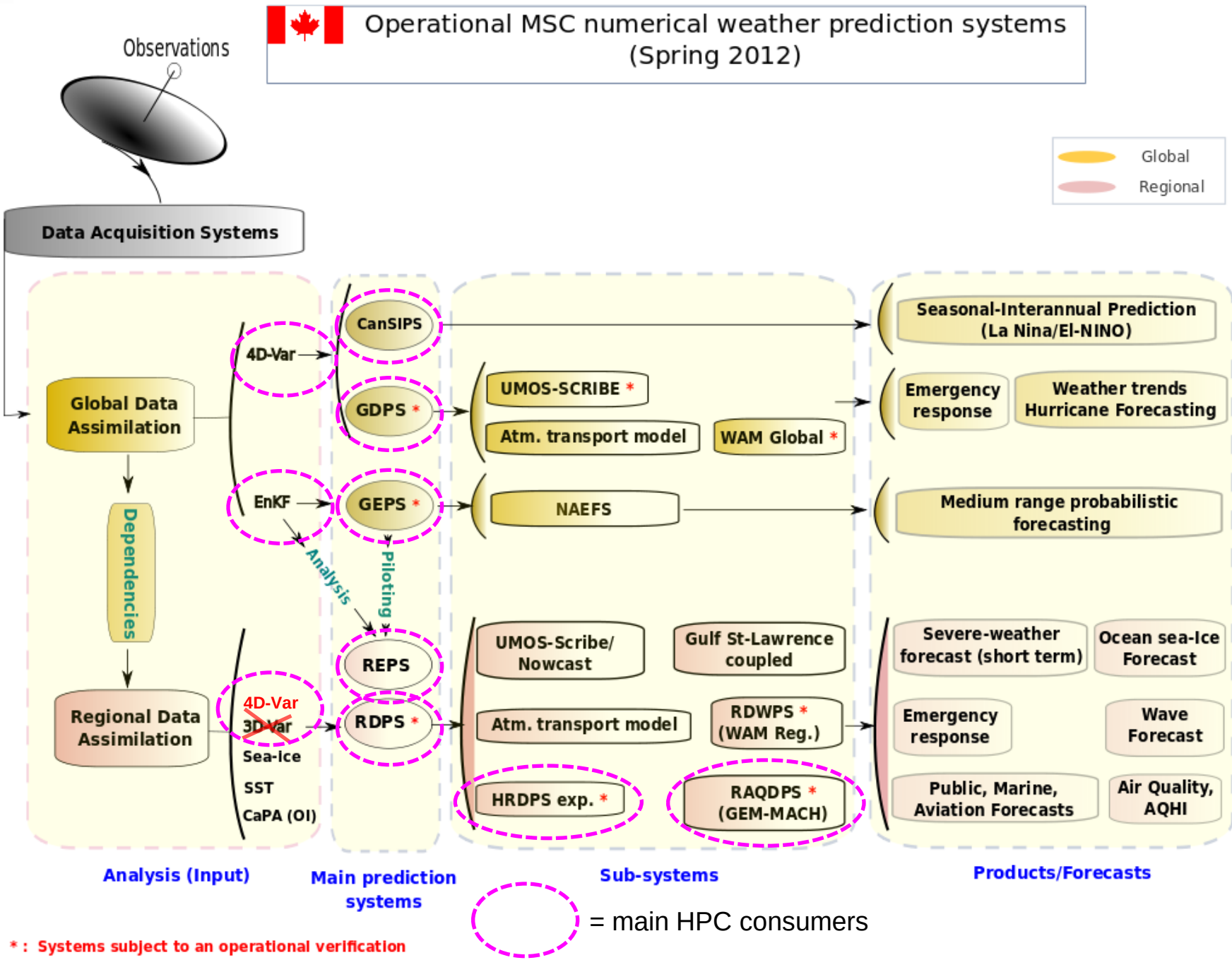
- Short-range deterministic forecasts (GEMREG-15km: was operational till 2010)
- Regional climate simulations
- Regional seasonal forecasts

Operational MSC numerical weather prediction systems (Spring 2012)

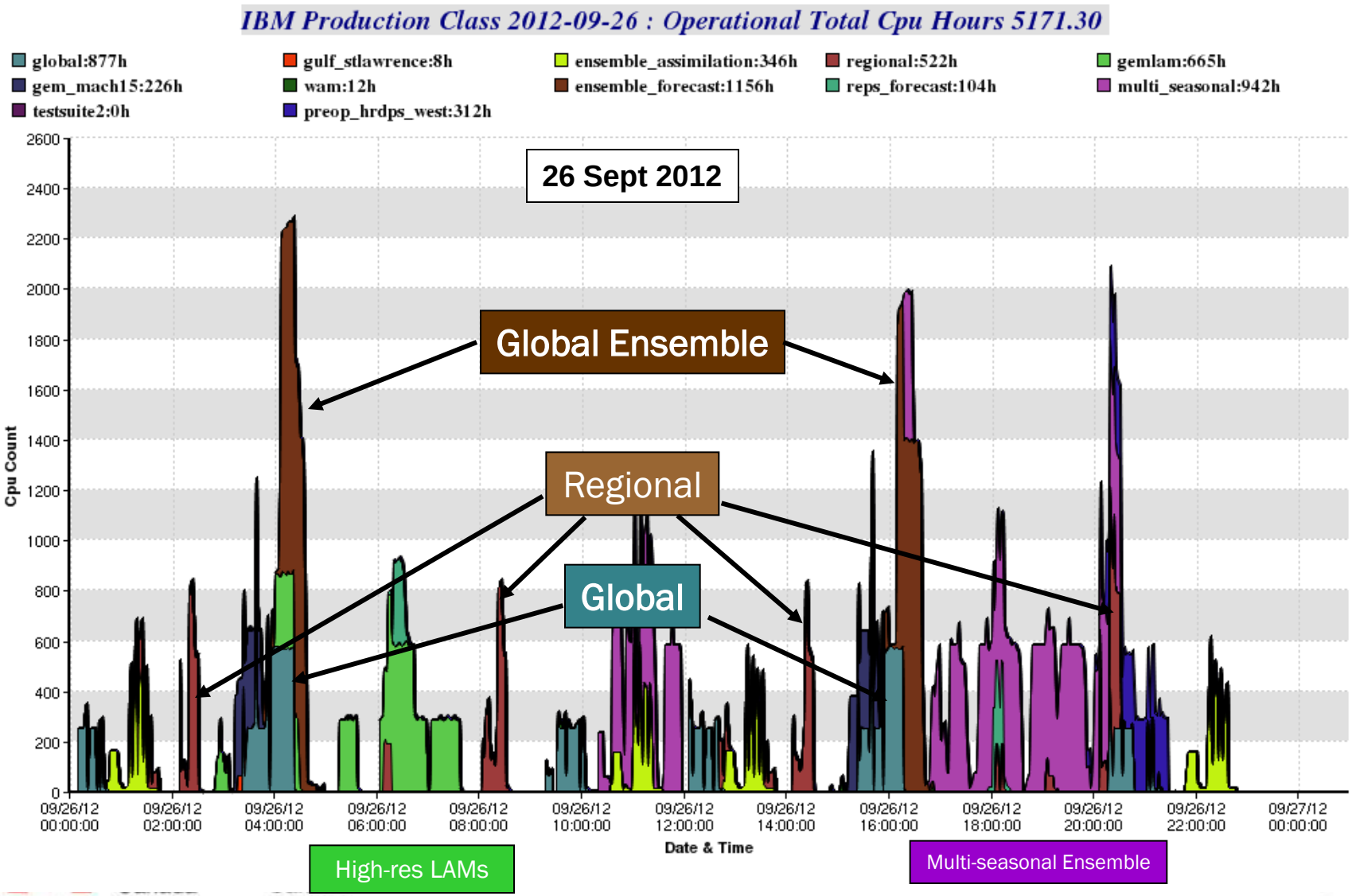


* : Systems subject to an operational verification

Operational MSC numerical weather prediction systems (Spring 2012)



Computational load of CMC NWP operational systems



Recent NWP operational implementations requiring significant HPC power

In 2011

For the first time in operation:

- Regional EPS (20 members; 33 km; North America)
- CanSIPS: Global coupled multi-seasonal prediction system (20 members; T63)
- Regional coupled prediction system for the Gulf of ST. Lawrence (15 km)

Recent NWP operational implementations requiring significant HPC power

In 2011

For the first time in operation:

- Regional EPS (20 members; 33 km; North America)
- CanSIPS: Global coupled multi-seasonal prediction system (20 members; T63)
- Regional coupled prediction system for the Gulf of ST. Lawrence (15 km)

Upgrade to existing systems:

- Global EPS
  - Assimilation (EnKF): From 96 to 192 members
  - Forecast: From 100 km to 66 km; from 28 to 48 levels; still 20 members
- Global Deterministic (33 km):
  - Forecast: modified deep convection to improve forecasting of tropical cyclones
  - Assimilation: 2 x more data assimilated; new SST analysis
- Regional Air Quality (15 km) : Updated emissions

Blue = significant computational load increase

Upcoming NWP operational implementations requiring significant HPC power

In 2012

This week !

- Regional (North America) prediction system: 4D-VAR; from 15 to 10 km;
- Regional (North America) Air Quality prediction system: from 15 to 10 km;
- Regional coupled prediction system (Gulf of ST. Lawrence): from 15 to 10 km;
- LAM 2.5 km:
  - Operational implementation of the West domain 2 x day;
  - Other 4 domains still experimental

Upcoming NWP operational implementations requiring significant HPC power

In 2012

This week !

- Regional (North America) prediction system: 4D-VAR; from 15 to 10 km;
- Regional (North America) Air Quality prediction system: from 15 to 10 km;
- Regional coupled prediction system (Gulf of ST. Lawrence): from 15 to 10 km;
- LAM 2.5 km:
  - Operational implementation of the West domain 2 x day;
  - Other 4 domains still experimental

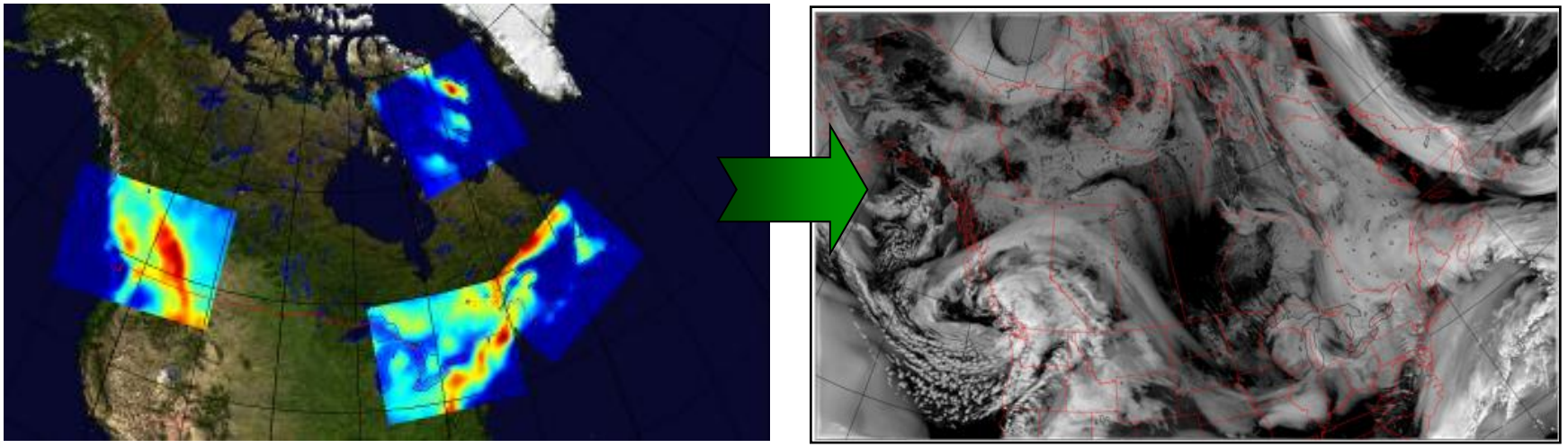
Before end of 2012 or in January 2013

- Global Deterministic :
  - from 33 km to 25 km;
  - Increased 4D-Var inner loop resolution (240x120 to 400x200)
- Global EPS :
  - Analysis (EnKF): Multi-scale (3x more assimilated data); From 100 km to 66 km;
- Regional EPS (20 members; North America): from 33 km to 15 km

Upcoming NWP operational implementations requiring significant HPC power

In 2013

- National LAM 2.5 km domain, 2 x day; replacing the LAM 2.5 windows



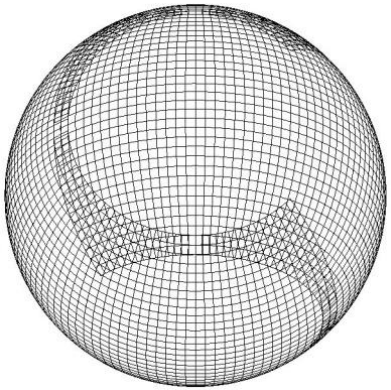
- Global Deterministic:
  - **Yin-Yang grid**
  - **from 25 km to 15 km**;
  - **EN-VAR** assimilation system replacing the 4D-VAR (no more TL/ADJ !)
- Regional (Arctic) Ice Prediction System (RIPS) (5 km) – Experimental
- Global Ocean (1/4 deg) - Experimental

Future evolution of CMC atmospheric models

	End of 2012	2016	2020
Global	Det. 25 km EPS 66 km	Det. 10 km EPS 35 km	Det. 10 km EPS 20 km
Regional	Det. 10 km EPS 15 km	Det. 2.5 km EPS 10 km	Det. 1.5 km EPS 10 km
Local	Det. *2.5 km*	EPS 2.5 km	EPS 1.5 km
Urban	*Det. 250-> 5m*	Det. 3m EPS 5m	Det. 1m EPS 5m

Future: new modeling approaches

- Yin-Yang grid
  - Very good scaling; no pole problem
  - **Operational in 2013 at 15 km resolution** on 75 P7 nodes (2400 cores)
    => **Global 240-h forecast in about 1h**



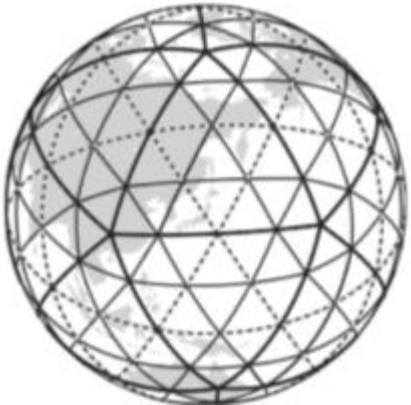
- Each piece is regular Lat/Lon grid

- Global forecast obtained by 2-way coupling of 2 LAM models

- Coupling done simultaneously at the solver level of both grids
=> no blending/relaxation of the two solutions needed

More info on Yin-Yang during next talk (Vivian Lee)

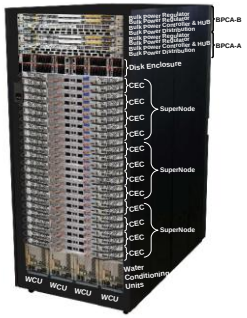
- Icosahedral grid
  - Scaling even better than Yin-Yang grid

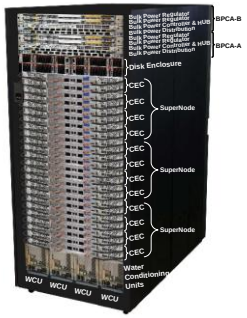


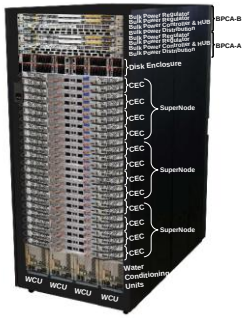
Environment Canada
New IBM Power7 Supercomputer

With permission from IBM

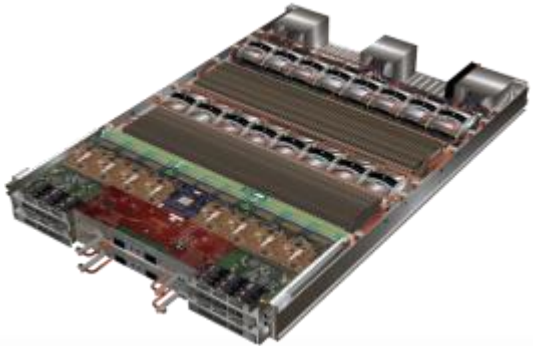
Environment Canada
New IBM P7 Supercomputer







- 2 clusters; 8192 cores each;
- ~ 1/2 PFlops peak total
- System was installed off-site
- Migration of operational jobs took a year to complete
- Fully operational since early May 2012



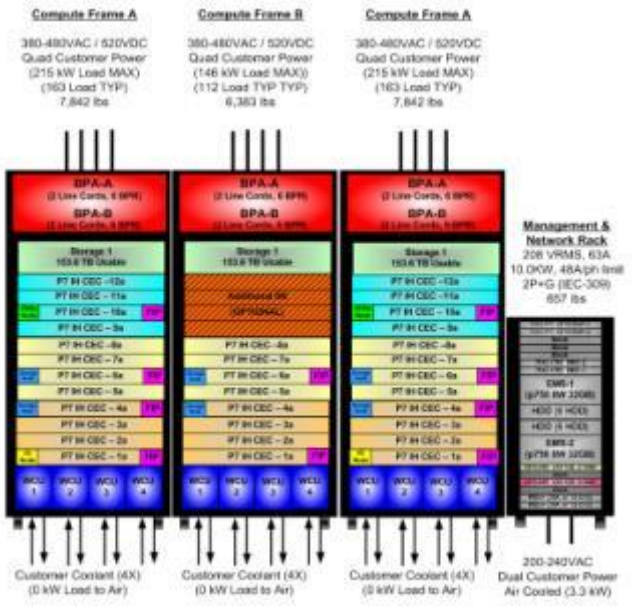
Environment Canada
New Supercomputer

- IBM Power7-based system
  - P775 units
  - Water cooled
  - **2500 kg** per cabinet !

- 2 clusters, each comprised of
  - 8 “super nodes” == 256 nodes (total, not all compute nodes)
  - 1 node = 32 cores
  - Total 8192 cores
  - 32.8 TB RAM (4GB/core)
  - 380 TB of usable GPFS shared storage per cluster
  - **242 TFlops peak (each cluster)**

Power7 installation timeline

- Equipment delivered late October 2011
- Installation
  - Operating system, storage, early-birds: Nov-Dec.
  - Parallel testing, stability improvements: Jan-April.
- Apr 11th: SP4 installation
- May 2nd: first operational products delivered by the P7
- May 11th: production ends on P5 system
- May 28th: P5 system powerdown
- Aug 15th: SP5 installation



Power 7 Outstanding Issues

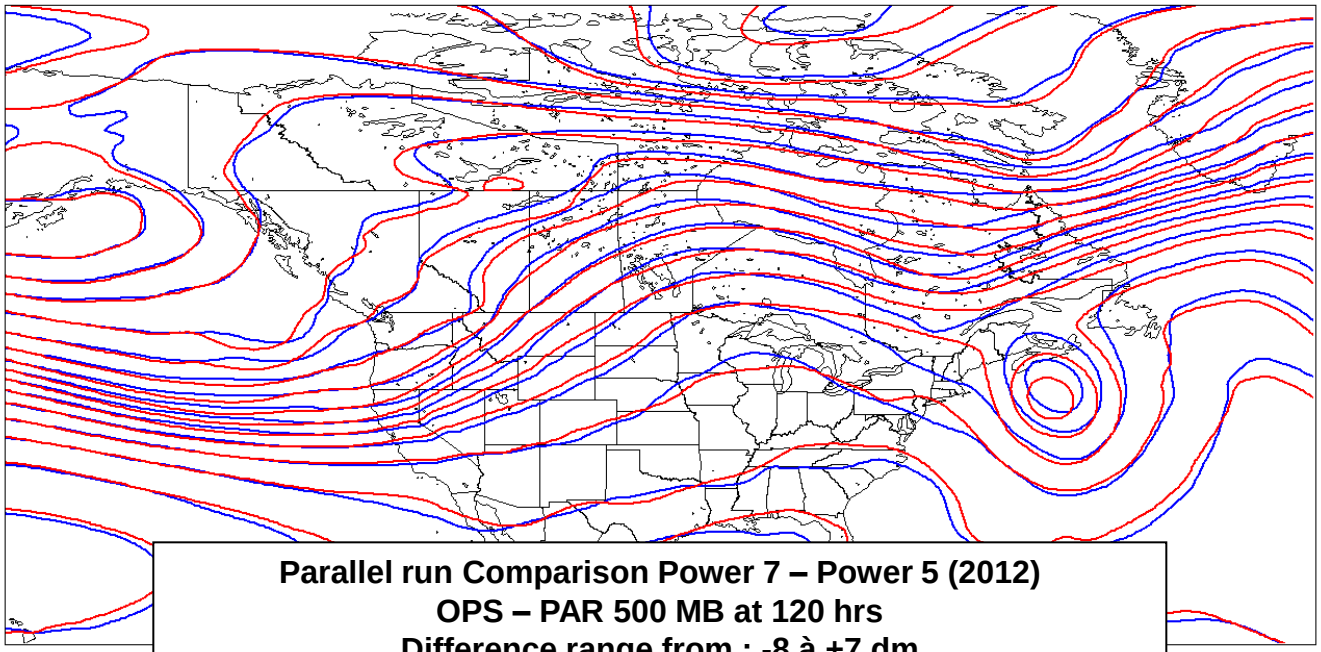
- HASN – High Availability for the Service Nodes is being deployed
  - Will improve overall cluster availability and ease of maintenance
  - Preemption dependent on HASN
    - Implementation of NFS over GPFS which will dramatically improve paging performance.
- xCat
  - New on P7 and being continuously improved by IBM
  - Initial rough edges have been smoothed
  - Issues remain but being addressed
- Site-wide LDAP interoperability issues have caused service interruption.

P5 to P7 model code migration

- Major project
  - Almost a year in the making but included a new R&D&O unified environment
  - Involved about 70 employees (not all full time!)
  - 26 coordination meetings at CMC, 26 for HPC user-reps nationwide
- Recompilation and testing of all the code
  - GEM in many configurations, Climate models, ocean models, etc
  - Support binaries and libraries
- The migration exercise was beneficial to model codes as it revealed coding weaknesses

P5 to P7 migration
-Validation-

- Results produced by P5 and P7 binaries were bit-for-bit identical for some models, but not for most



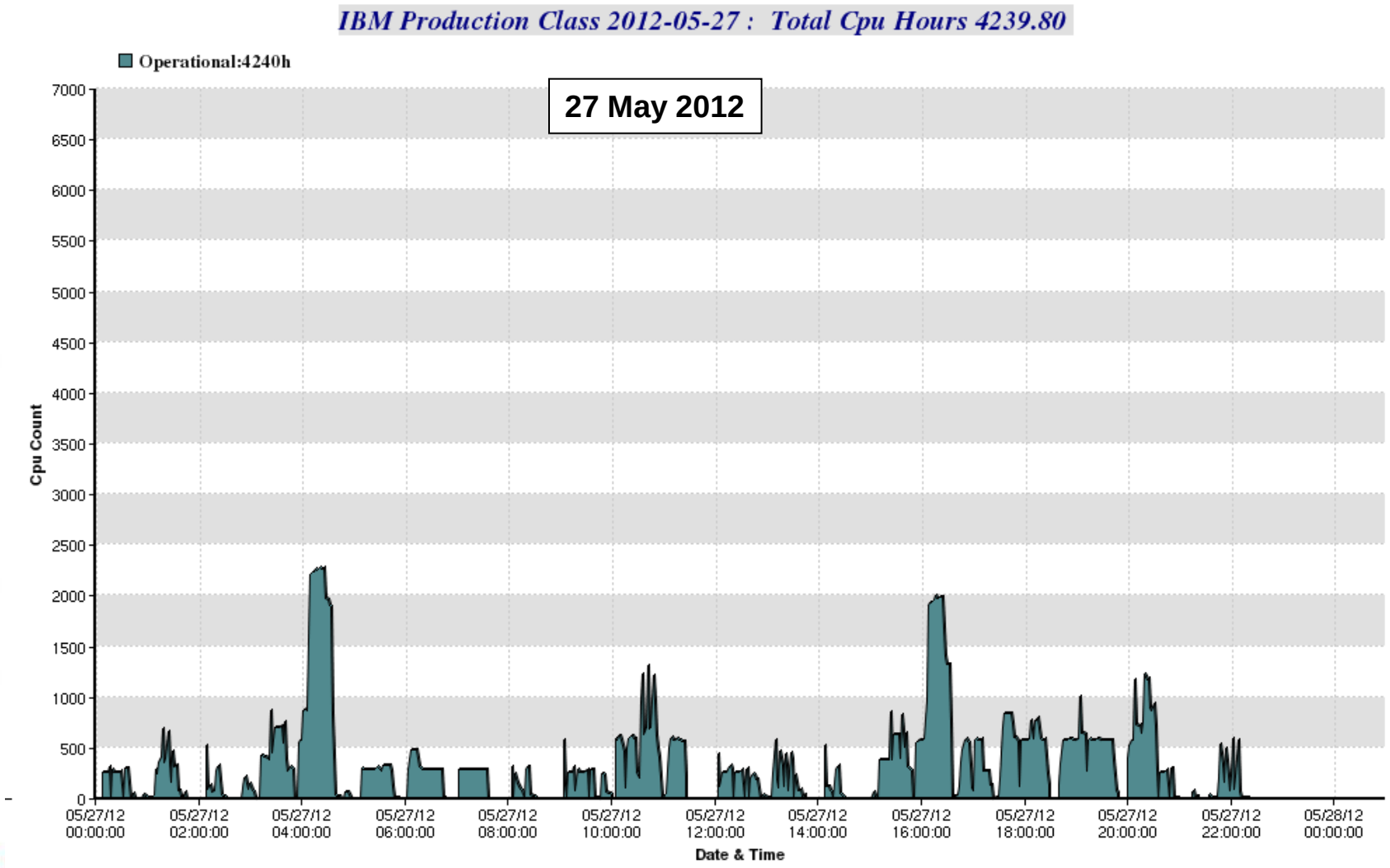
Parallel run Comparison Power 7 – Power 5 (2012)
OPS – PAR 500 MB at 120 hrs
Difference range from : -8 à +7 dm

- Statistical tests were performed as well as evaluation of parallel runs by expert operational meteorologists

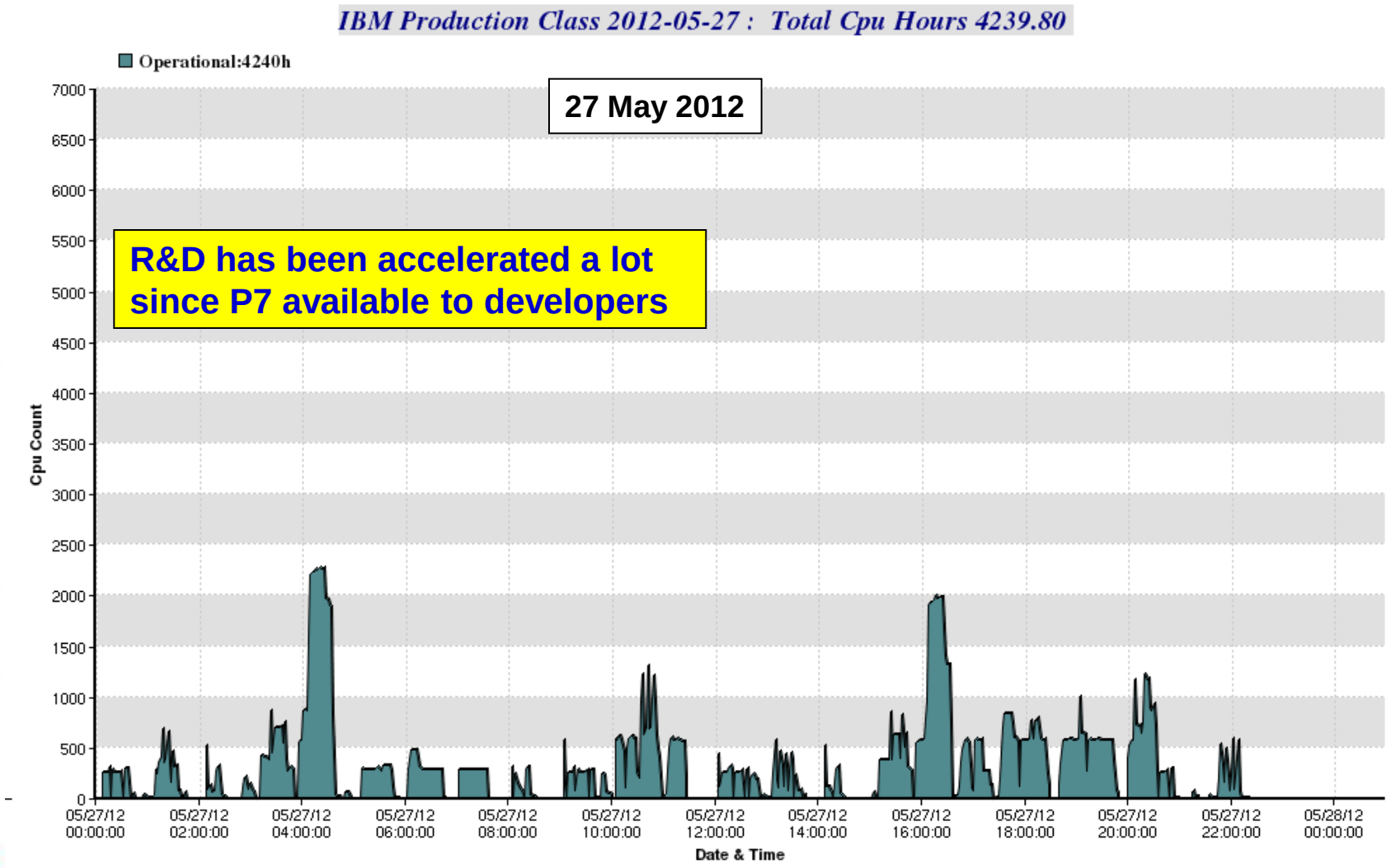
Environment Canada
Power 7 performance

- Overall performance
  - On average, performance gain per CPU about 2.7x compared to P5 CPU
- System performance
  - Pre-operational stability issues compared to migration to P4 and P5 which were more mature
  - We had to go through 4 major service packs (SP2 to SP5), with some significant down times during the service pack installations with no impacts on operations
  - Delocalisation of the supercomputer had no impacts (very good data links)

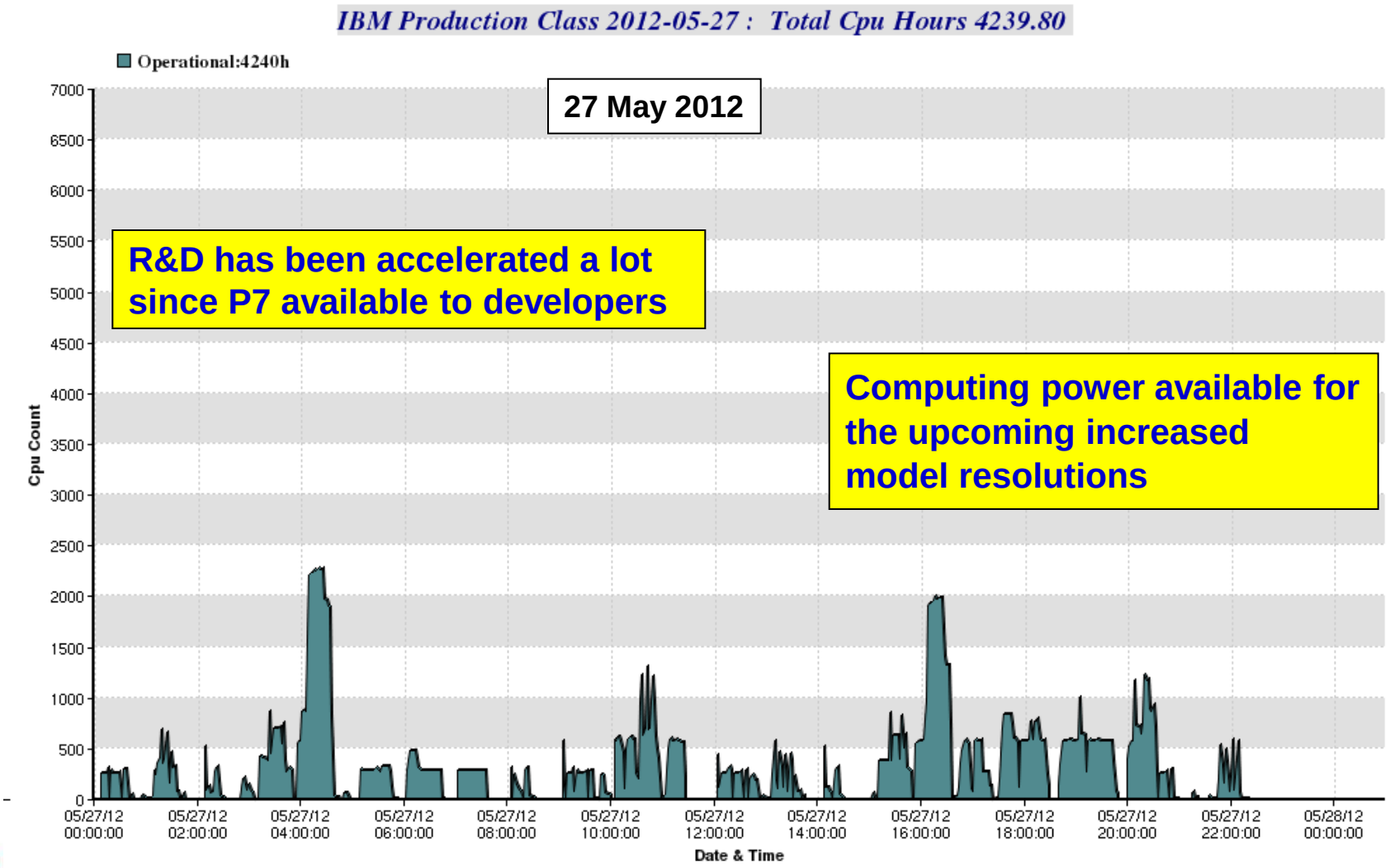
Load of CMC NWP operational systems on the new operational IBM P7 cluster



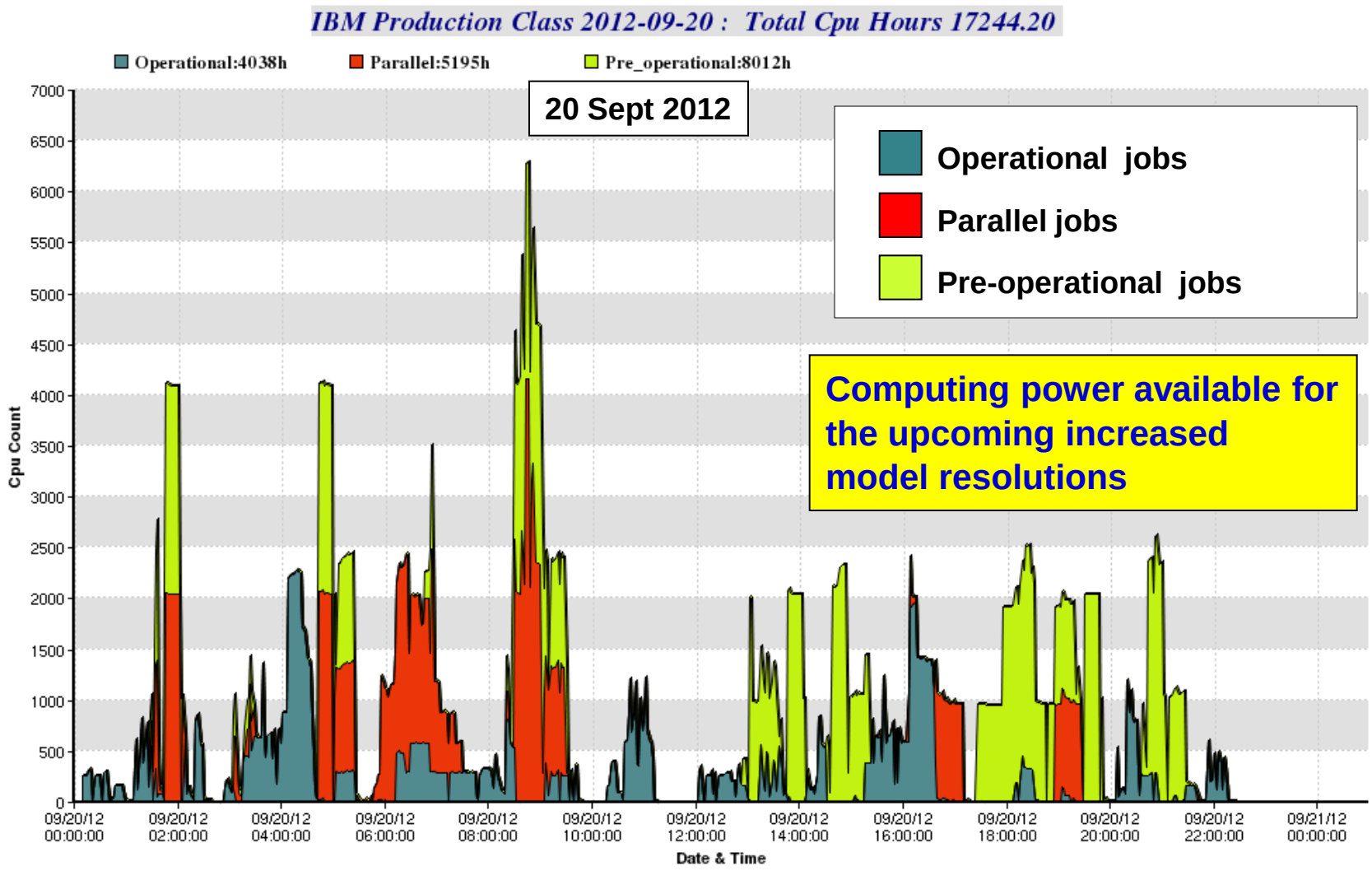
Load of CMC NWP operational systems on the new operational IBM P7 cluster



Load of CMC NWP operational systems on the new operational IBM P7 cluster

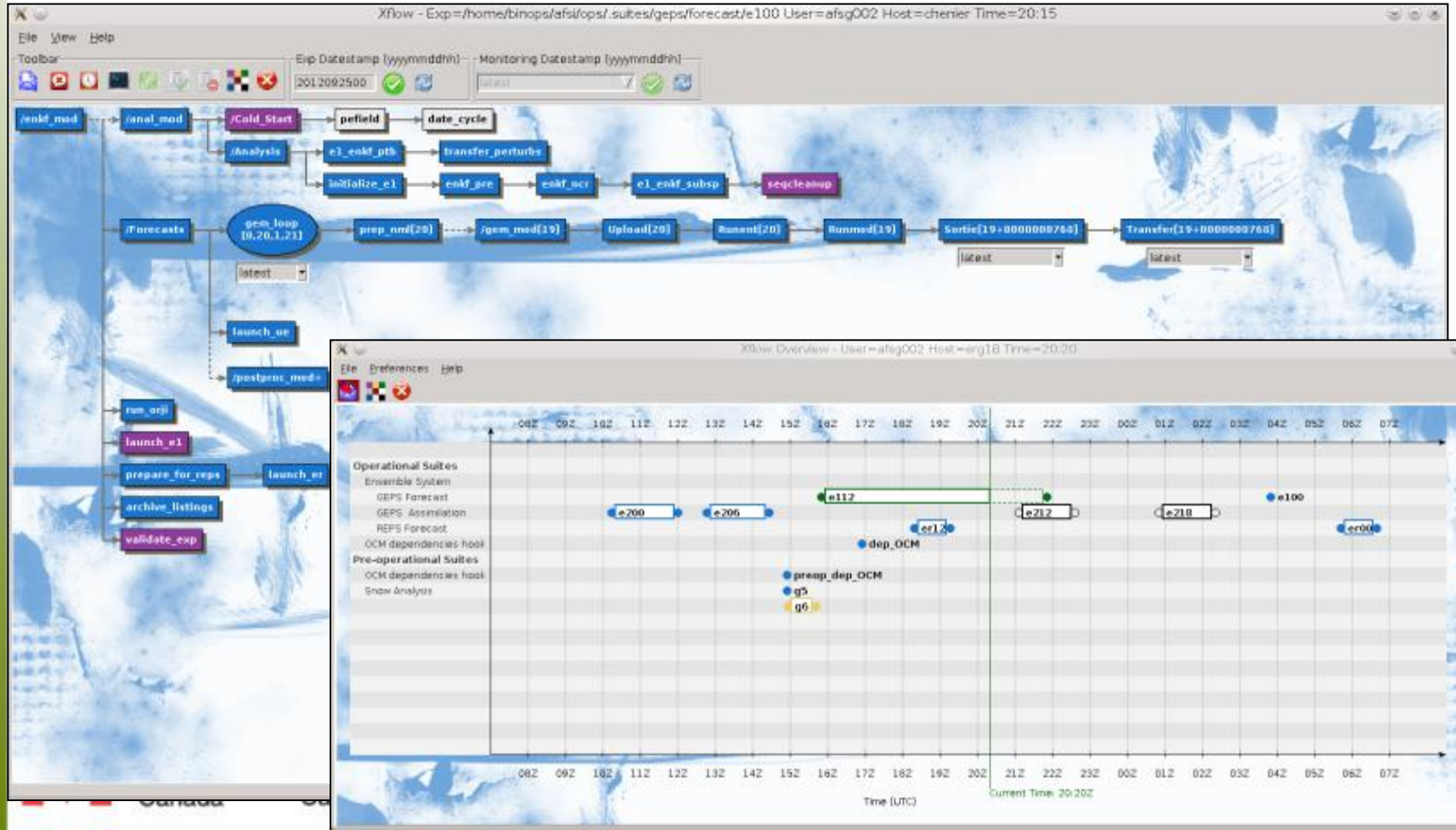


Parallel / pre-op /operational / systems load on the new operational IBM P7 cluster



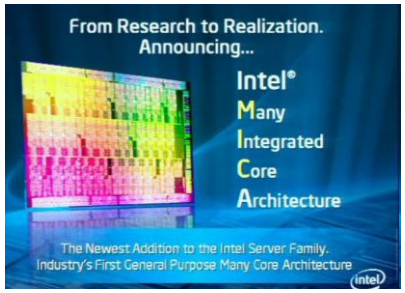
Common set of tools and coding standards for R&D&O

- New sequencer (Maestro) and coding standards (based on industry standards)
- Objective: reduce the high cost of tech transfer
- Challenge: convincing people to change their ways



Preparing post - IBM Power 7 era

- We know our HPC users' needs for the next ten years
  - Modeled physical processes will continue to increase in complexity
  - Model resolution will continue to increase, and aggressively
  - Ensemble paradigm will be stronger than ever
  - The number and the complexity of coupled environmental modeling systems will continue to grow
- Model optimization, emerging technologies
  - Many Core Integrated (MIC) or GPU or Hybrid ??

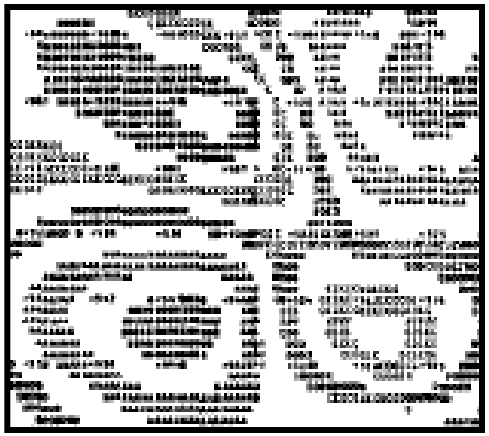


Now that the IBM Power 7 is being filled up quickly, the time has come for preparing the post Power 7 era

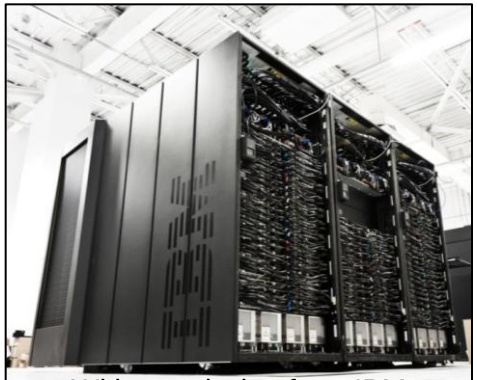
Thank you !

Bendix G20, Montreal during the 60's

500 mb geopotential forecast

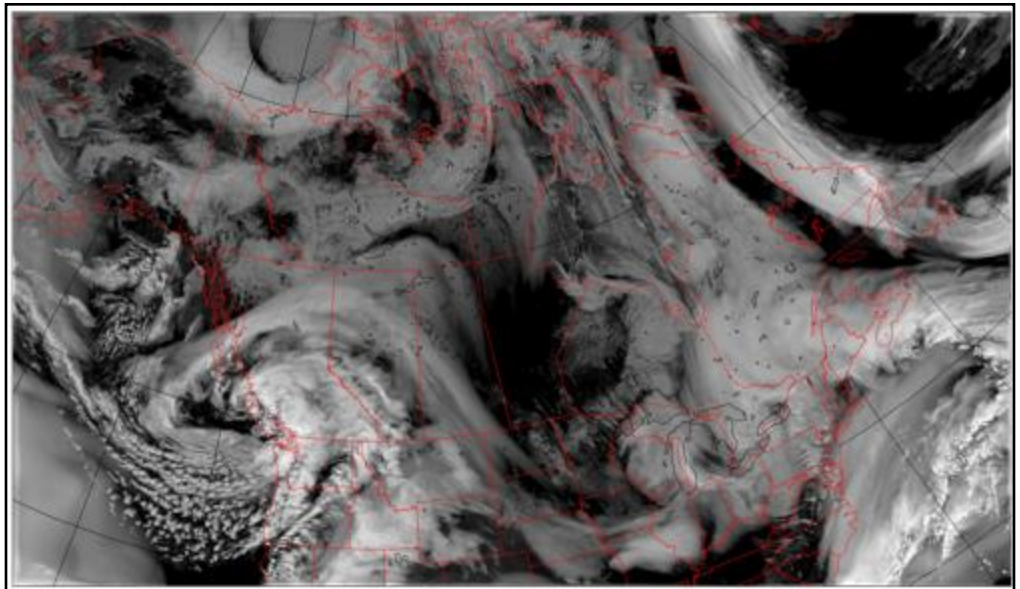


2012: IBM Power 7



With permission from IBM

National coast-to-coast-to-coast LAM 2.5 km



Environment Canada Environnement Canada