NCEP Applications -- HPC Performance and Strategies

Mark Iredell
software team lead
USDOC/NOAA/NWS/NCEP/EMC

ECMWF Workshop HPC in Meteorology
1 October 2012

Motivation and Outline

- **Challenges in porting NCEP applications to WCOSS and future operational systems**
- **Major scientific challenges are driving major HPC challenges**
- **Our strategy:**
  NOAA Environmental Modeling System
- **Challenges in running a fully interconnected system in operations**

WCOSS Contract
System Characteristics (per site)

	Life Cycle	Arch	OS	Average Capability	Average Capacity	Nbr of compute / batch cores	TeraFLOP
Bridge System	Oct 2012 - Sep 2013	Power6	AIX 5.3	1.0X	1.0X	5,314	73.9 TF
WCOSS Phase 1	Plan accept Dec 2012 – FOC Aug 2013	iDataPlex	Linux (RHEL)	2.0X over Bridge P6	2.3X over Bridge P6	10,048	208 TF
WCOSS Phase 2	Plan accept Dec 2014 – FOC Jul 2015	iDataPlex	Linux (RHEL)	*1.9X over Phase 1	*4.4X over Phase 1	*44,400	*920 TF

***Estimated values; assumes FY13 NWS Reallocation Budget**

WCOSS Contract
System Characteristics (per site) (cont.)

	Life Cycle	Operational Use Time Requirement	Storage Useable	Power KW	Floor Space
Bridge System	Oct 2012 - Sep 2013	99%	0.80 PB	689 KW	3,100 SF
WCOSS Phase 1	Plan accept Dec 2012 – FOC Aug 2013	99.9%	2.59 PB	469 KW	4,060 SF
WCOSS Phase 2	Plan accept Dec 2014 – FOC Jul 2015	99.9%	*7.2PB	*1050 KW	4,060 SF

*** Estimated values; assumes FY13 NWS Reallocation Budget**

Short-term porting challenges

NCEP will port all applications from CCS to WCOSS over the next few months. Below is a summary of the major differences that must be addressed.

- 16 processors per node rather than 32
- 32 GB memory per node rather 128 GB
- Little endian rather than big endian
- Different libraries
- Different compiler
- Different tools
- Different batch system
- Different scheduling system

Long-term HPC challenges

Major scientific challenges driving major HPC challenges:

- Coupling large ensembles with interdependencies
- Coupling atmosphere and ocean and ice and waves
- Coupling ionosphere model with atmosphere model
- Coupling global scale with higher and higher resolution
- Massive output requiring parallel I/O
- Changing HPC architectures

NCEP is looking toward the NEMS framework to address these challenges.

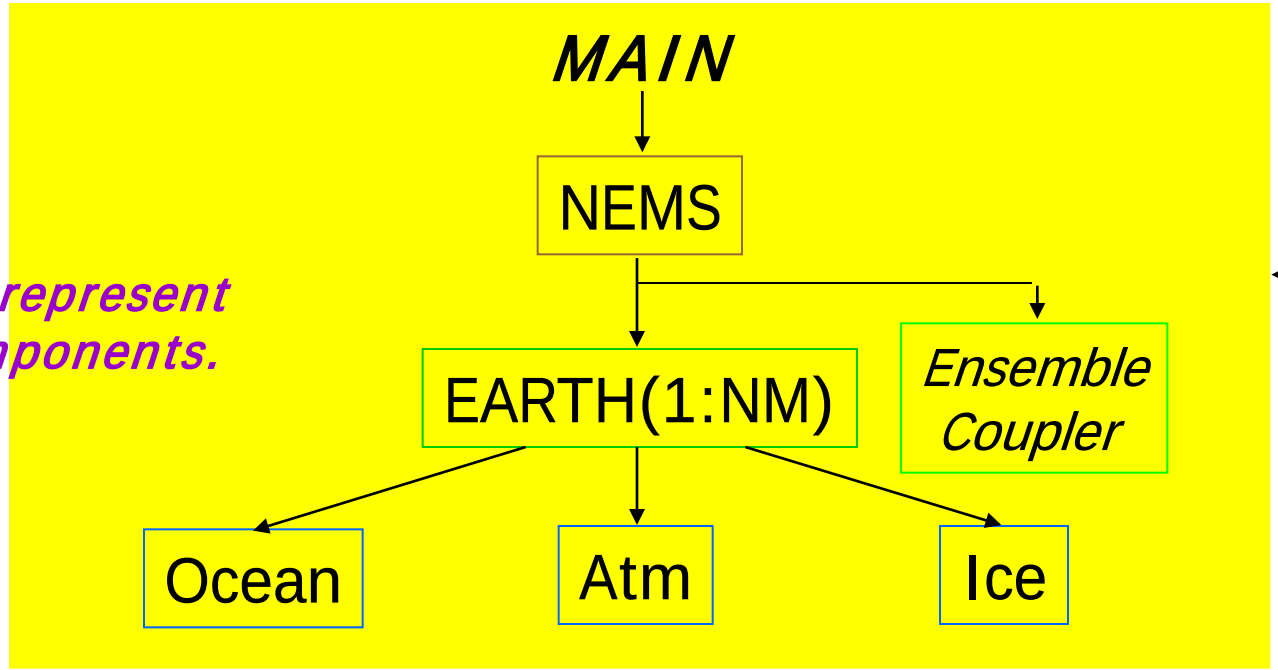
What is NEMS

- NOAA Environmental Modeling System
- A superstructure framework to couple environmental models and provide High Performance Computing tools for parallel coupling and parallel I/O for models
- Uses the Earth System Modeling Framework (ESMF) and National Unified Operational Prediction Capability (NUOPC)
- Couples ensemble members together
- Couples global scale with mesoscale and microscale
- Couples atmosphere with ocean and wave and ice and land and ecosystems and space
- Couples ionosphere with atmosphere
- Accommodates future HPC architectures

Where is NEMS now

- North American Model (NMM-B)
  implemented operationally in 2011
- NEMS GFS Aerosol Component
  implemented operationally in 2012
- Global Forecast System (global spectral model)
  targeted to be implemented operationally in 2013
- Whole Atmosphere Model (WAM) up to 600 km
- Finite volume icosahedral model (FIM)
  installed and maintained by NOAA ESRL
- NCEP Global Ensemble Forecast System (GEFS)
- Multigenerational moving nests in mesoscale component
- Future ocean models HYCOM and MOM and wave model and sea ice model and couplers
- Future ionospheric model coupled to WAM
- Ported to CCS, Gaea, Zeus, and WCOSS

NEMS Framework



All boxes represent ESMF components.

NEMS super-structure

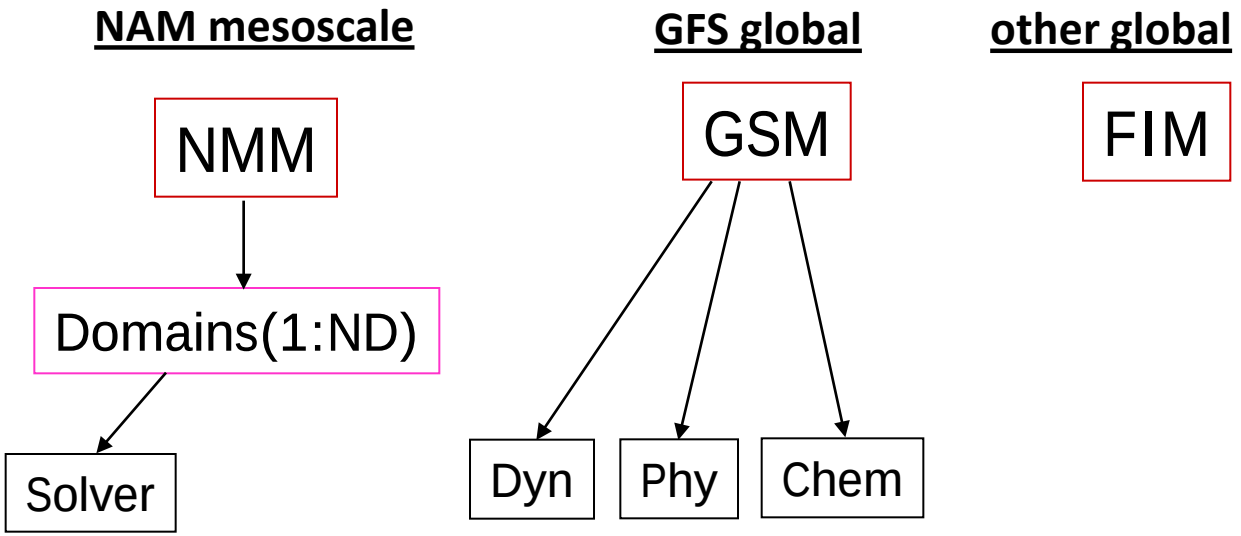
Wrt

nemsio, ...

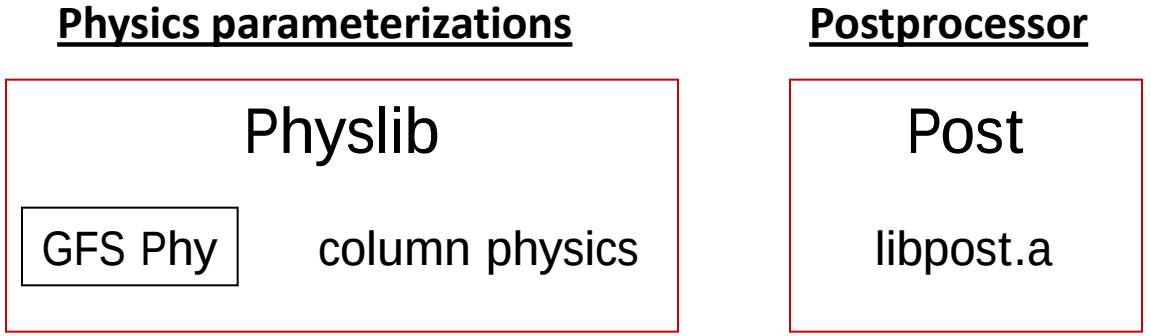
NEMS library

Components under NEMS

Model components:



Other components:



Two HPC challenges

Running all components in a single execution of NEMS presents operational HPC challenges…

- **Fault tolerance:** If one component fails for whatever reason, some soft recovery should be possible rather than a hard crash of the whole system
- **Validation:** Components should be able to be validated by themselves as well as when they run as part of the whole system

Fault tolerance

- ESMF 5.3 has fault tolerance prototype
  - only works for MPMD as yet
  - only works if inter-component MPI not used
- NEMS cannot take advantage yet
  - but may be able to in a future ESMF version
- NCEP would like a robust fault tolerance
  - run “stunted” ensemble if one member fails
  - apply a backup strategy if a component fails

Validation (part one)

System tests and Unit tests

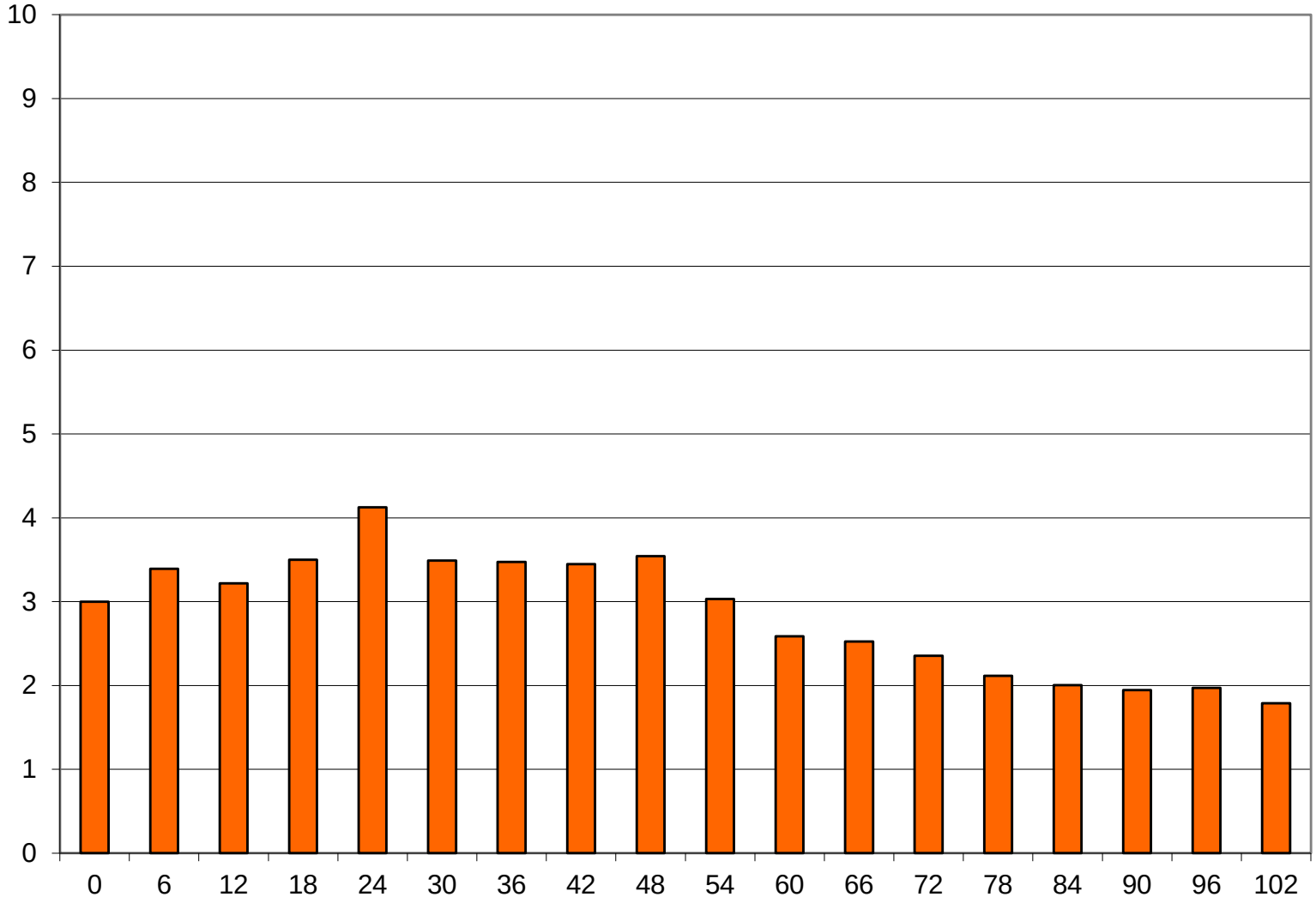
- NEMS regression tests
  - short list that runs with each commit
  - long list that runs once a week
- Component unit tests
  - run for each component or algorithm
  - demonstrate behavior and ensure robustness

Validation (part two)

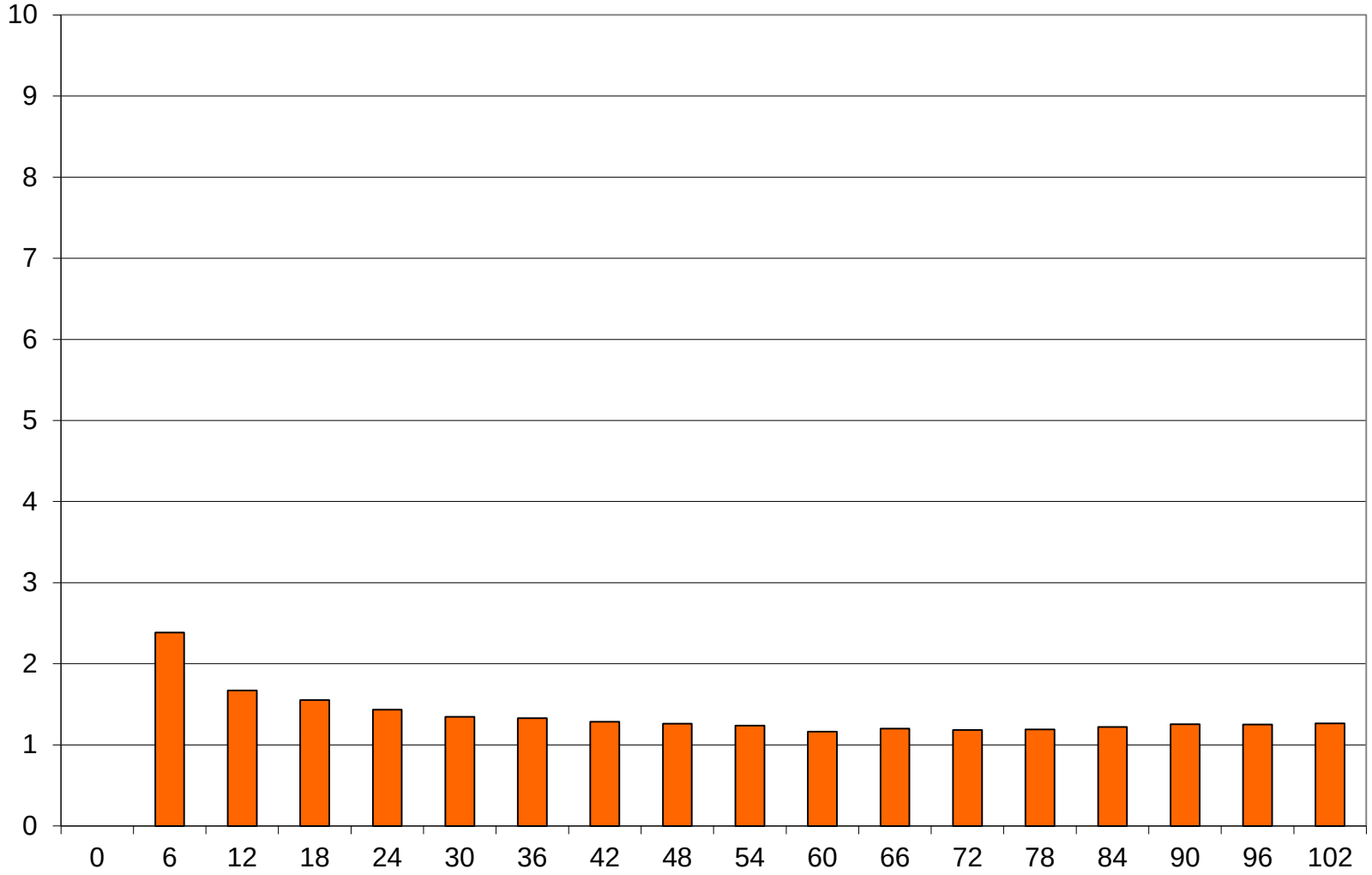
Functional Equivalence

- Regression test to count butterflies
  - developed for minor restructuring and porting
- Formalize functional equivalence
  - run control C, perturbed control P, and targeted improved or ported version T
  - perturbed control has input with last bit flipped
  - #butterflies = RMS(T-C)/RMS(P-C)

#butterflies: *GFS port to Zeus*



#butterflies: *NAM port to Gaea*



Summary and Conclusions

- NCEP transitioning to new WCOSS system
- NEMS addresses more interconnectedness
- NEMS shall readily accommodate and adapt to new HPC architectures
- More fault tolerance and validation required

Thank you. Questions?