Weather and Climate Operational Supercomputing System (WCOSS) - An operational perspective

Steven Earle

Lead Production Analyst

NWS/NCEP Central Operations

February 20, 2019

Outline

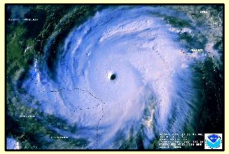
- Introduction to NCEP Central Operations and WCOSS
- History of supercomputing at NCEP
- Today's system - WCOSS
- Today's Computers -- WCOSS
- Research to operations - Challenges
- The future WCOSS

NWS National Centers for Environmental Prediction

Aviation Weather Center
Kansas City, MO

Space Weather Prediction Center
Boulder, CO

Storm Prediction Center
Norman, OK

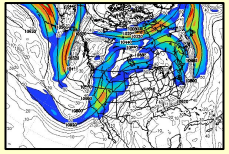


National Hurricane Center
Miami, FL

NCEP Central Operations
College Park, MD
(Supercomputers in Reston & Orlando)

Ocean Prediction Center
College Park, MD

Climate Prediction Center
College Park, MD



Environmental Modeling Center
College Park, MD

Weather Prediction Center
College Park, MD

Mission

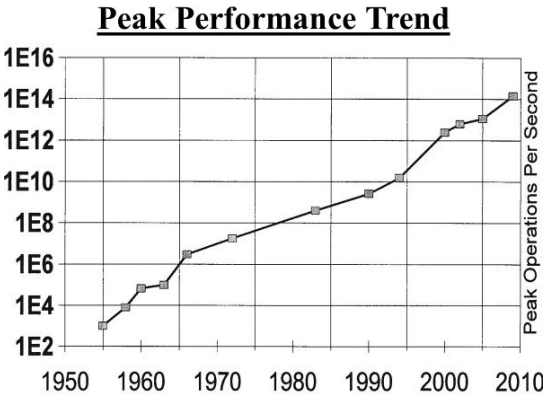
NCEP delivers national and global operational weather, water and climate products and services essential to protecting life, property and economic well-being.

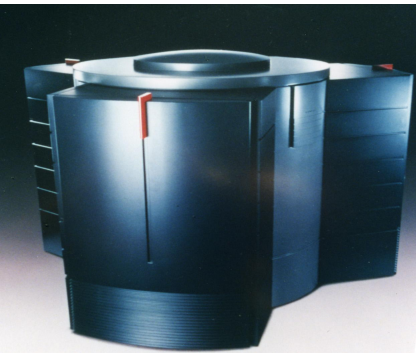
Vision

The trusted source for environmental predictions from the sun to the sea, when it matters most.

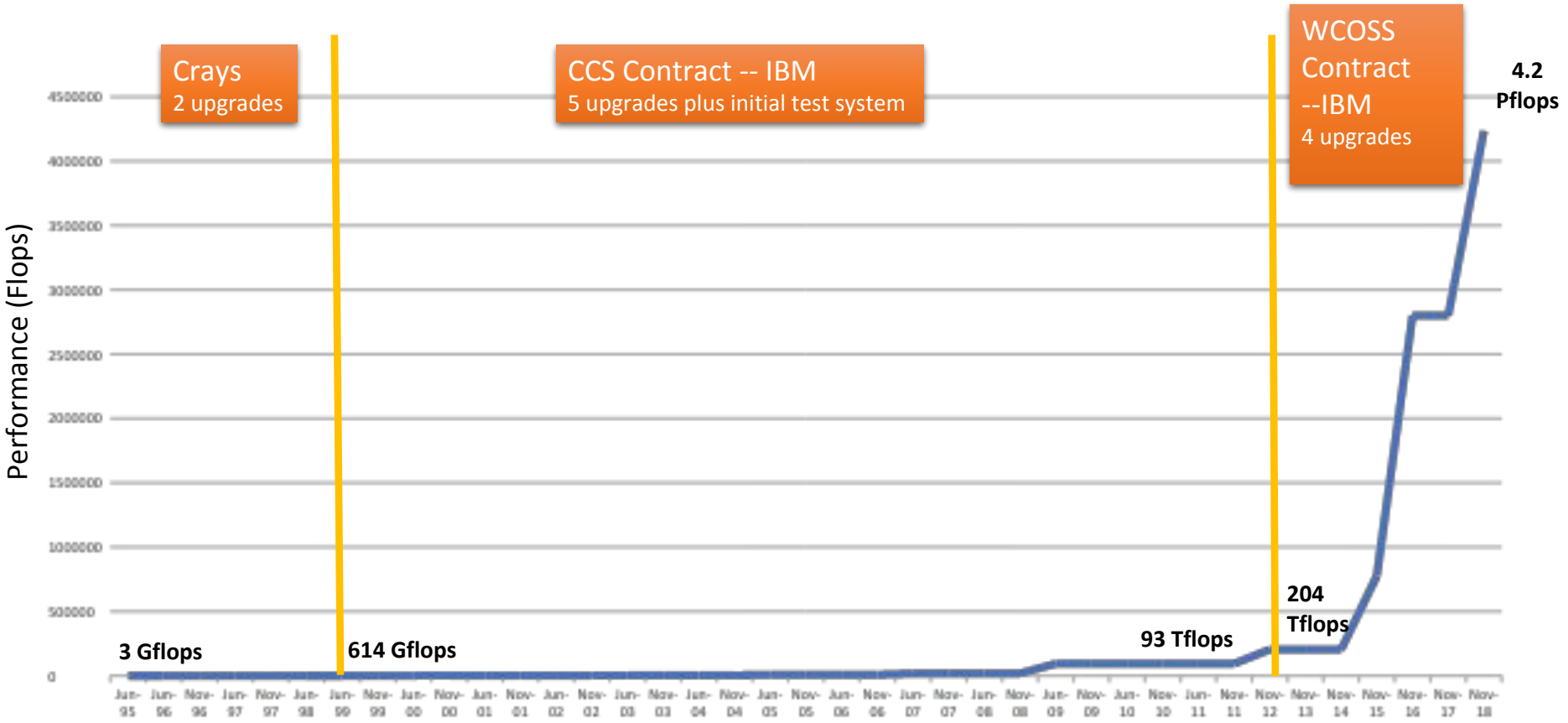
Supercomputing in NCEP

The National Weather Service purchased our first supercomputer in 1955 – the last of the IBM 701s



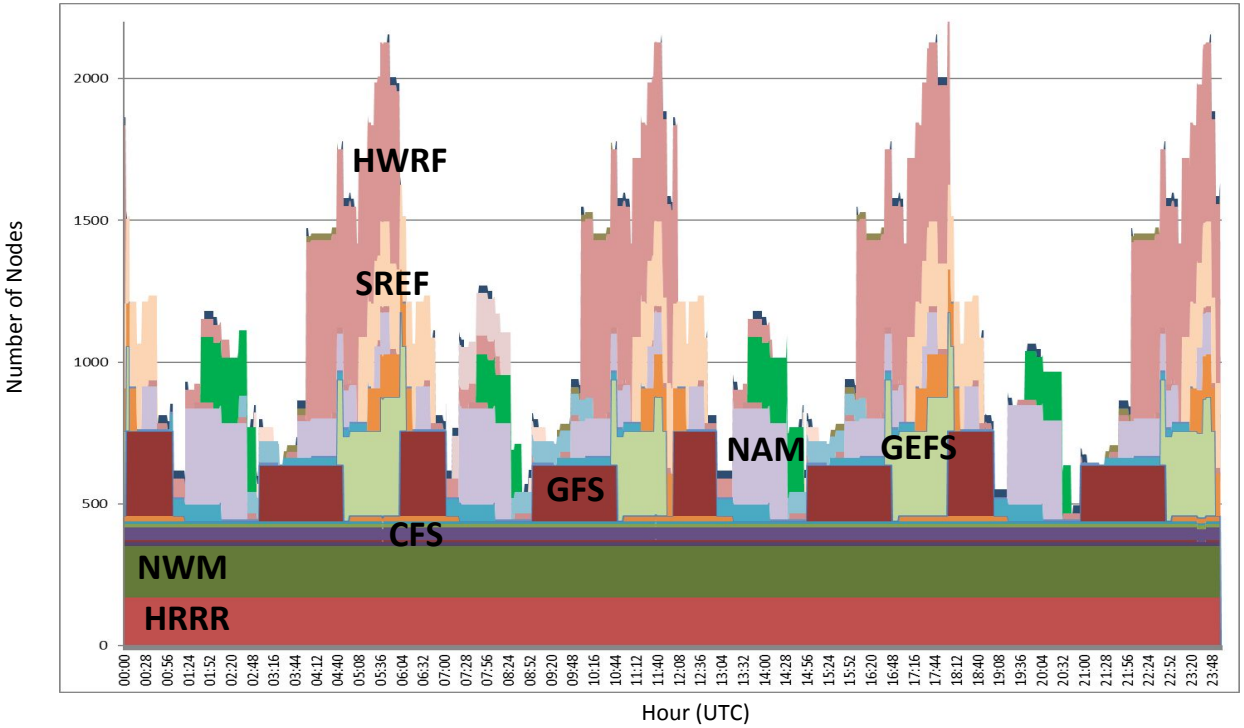


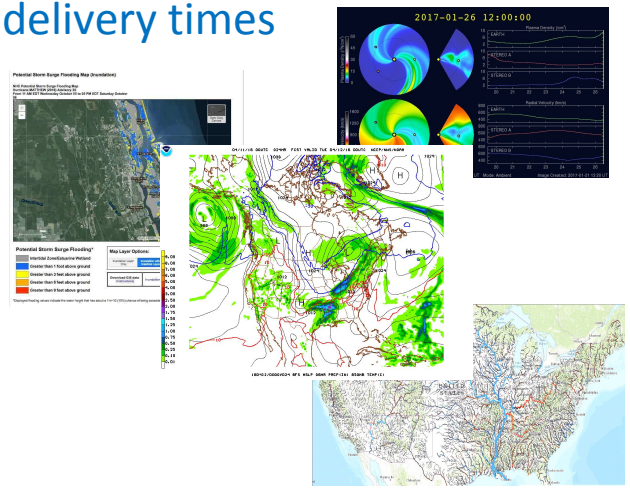
NWS Operational Computing Over the Last 20 Years



NCEP Production Suite (NPS)

- 30+ major models and associated applications from across NOAA – air, water and space
- 24x7 operation. Goal: routine, reliable and consistent product delivery times





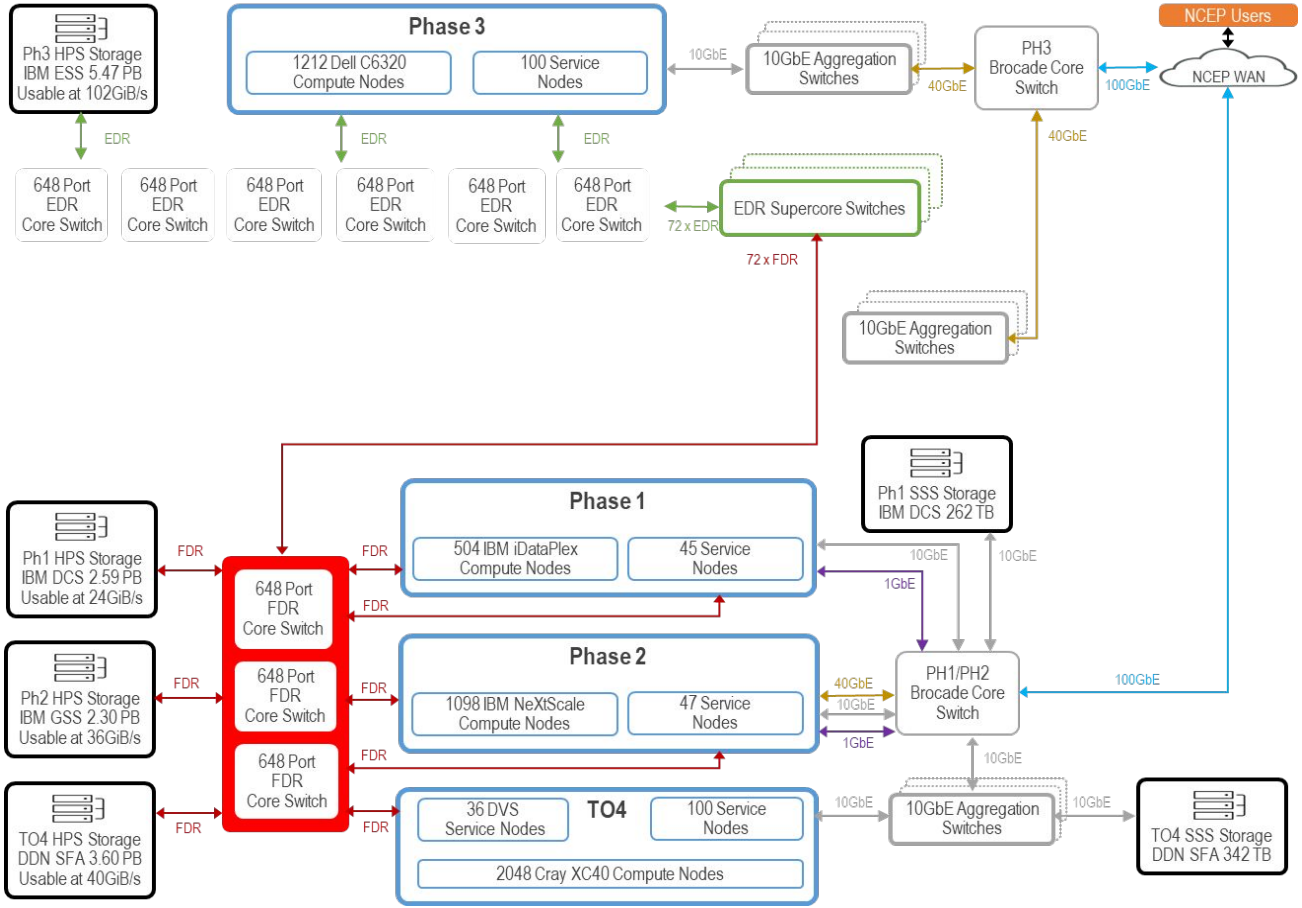
GFS – Global Forecast System
CFS – Climate Forecast System
NWM – National Water Model
HRRR – High Resolution Rapid Refresh
GEFS – Global Ensemble Forecast System
NAM – North American Model
SREF – Short-Range Ensemble Forecast
HWRF – Hurricane Model

Current Computing -- WCOSS

- 10 year contract awarded to IBM from 2011 through 2021. Includes initial delivery, three subsequent upgrades.
- Supplemental funding from Congress added more compute
- 2 identical clusters -- Orlando, FL and Reston, VA
- Currently 4.2 Pflops, 14 PB disk, 5260 nodes
- Heterogeneous system -- Combination of IBM iDataPlex, Cray XC40, and Dell PowerEdge hardware
  - Chips include Sandy Bridge, Ivy Bridge, Broadwell and Haswell
- Simultaneous Production and Development workload – 500+ users
- Produces 140 million products/day, distributes over 10 TB of guidance/day

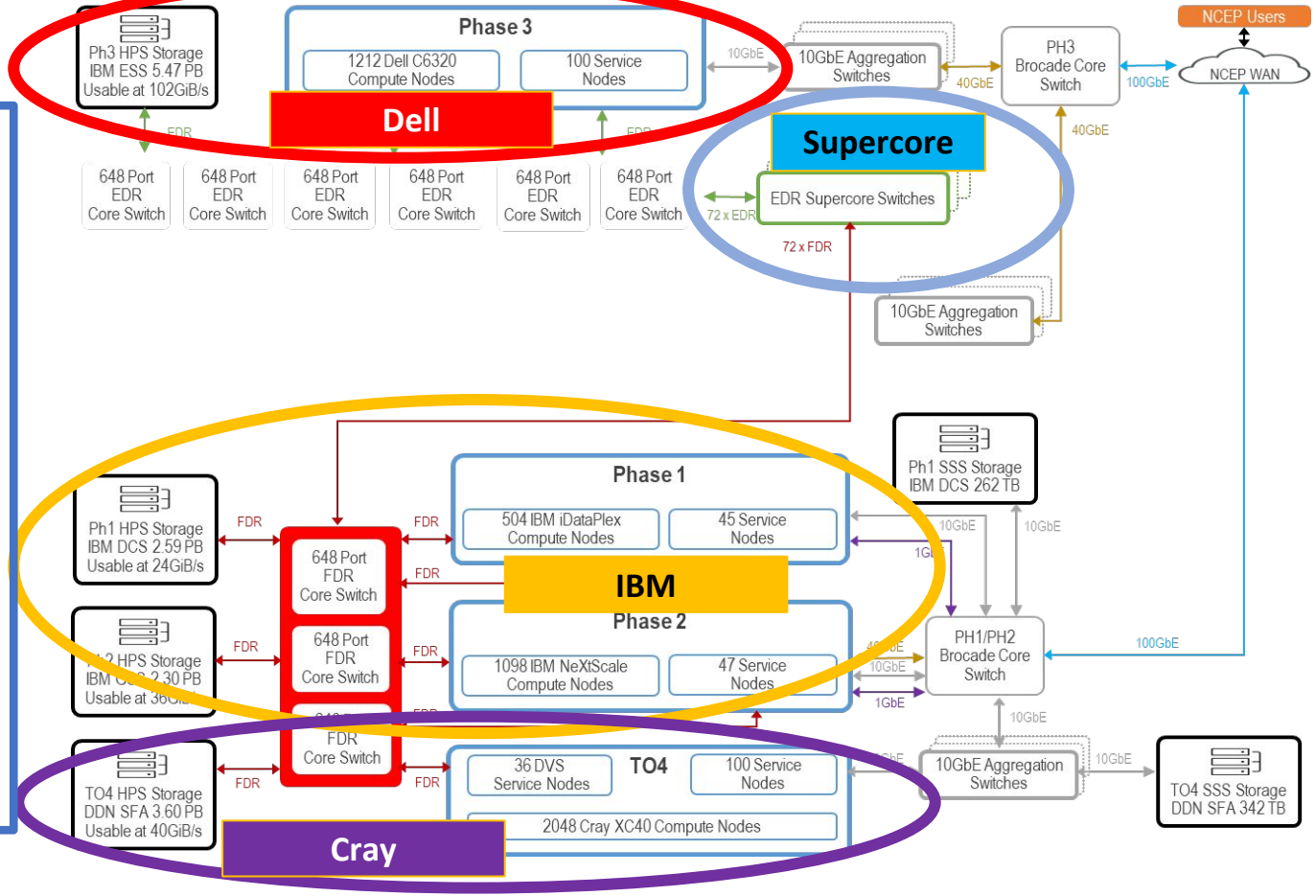
WCOSS Schematic

- 4948 compute nodes
- 312 service nodes
- Filesystem is GPFS (Spectrum Scale)
  - DCS
  - GSS
  - DDN
  - ESS
- Shared Storage to move data btw machines
- Mellanox and Infiniband interconnect
- Brocade Switches



WCOSS Schematic – Heterogeneous System

- 4948 compute nodes
- 312 service nodes
- Filesystem is GPFS (Spectrum Scale)
  - DCS
  - GSS
  - DDN
  - ESS
- Shared Storage to move data btw machines
- Mellanox and Infiniband interconnect
- Brocade Switches

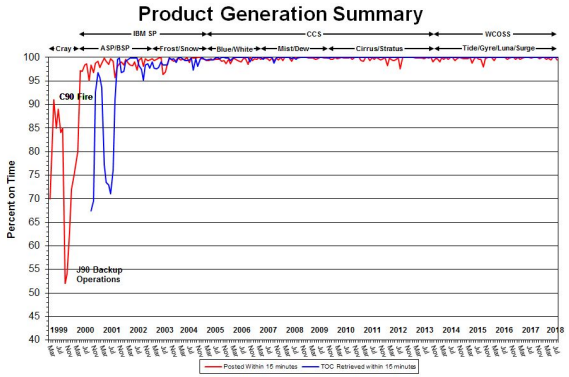


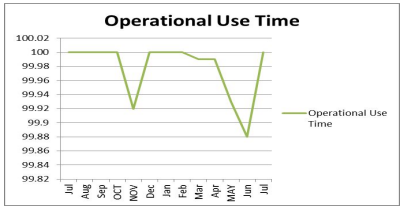
WCOSS -- Providing Operational Resiliency/Reliability

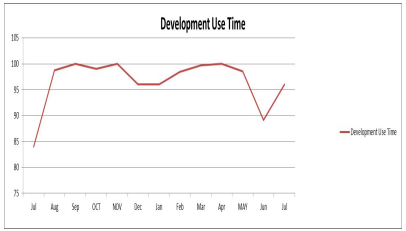
Performance Requirements

- Minimum 99.9% Operational Use Time
- Minimum 99.0% On-time Product Generation
- Minimum 99.0% Development Use Time
- Minimum 99.0% System Availability

- Penalties in contract for failing to meet metrics







Two Identical Systems – one Production and one Development

- Production locked down to NCO and select users
- Development machine is open to all users
- Switch between systems takes 15 minutes, but models bleed off for a few hours.
- Data mirrored between two systems – 40TB per day. Must have that for failovers and development work.
- For major maintenance activities, only one system is down. Production remains unaffected.

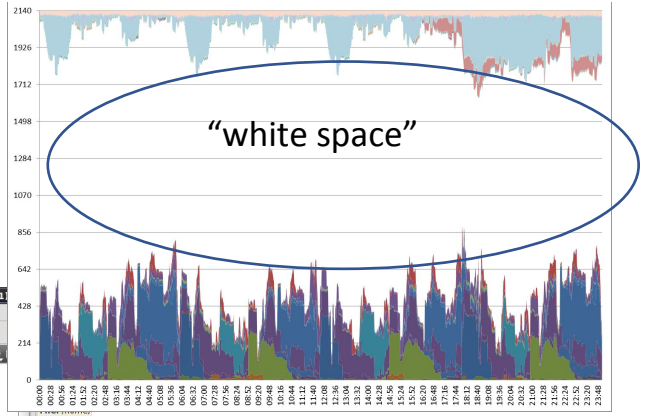
Three Tier Approach

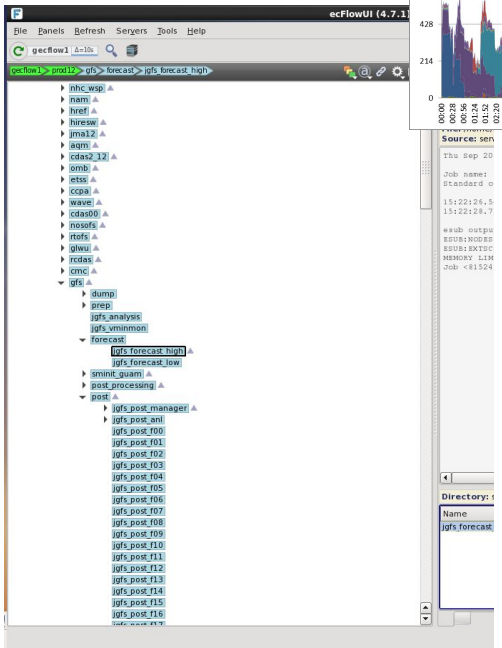
- **Tier 1 -** Operations Monitoring Branch
- **Tier 2 -** Implementation and Data Services Branch
- **Tier 3 -** Development organization

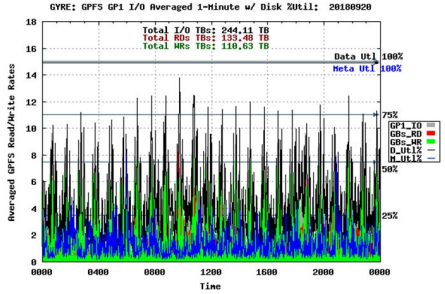
WCOSS – Workload Management

Managing Workload

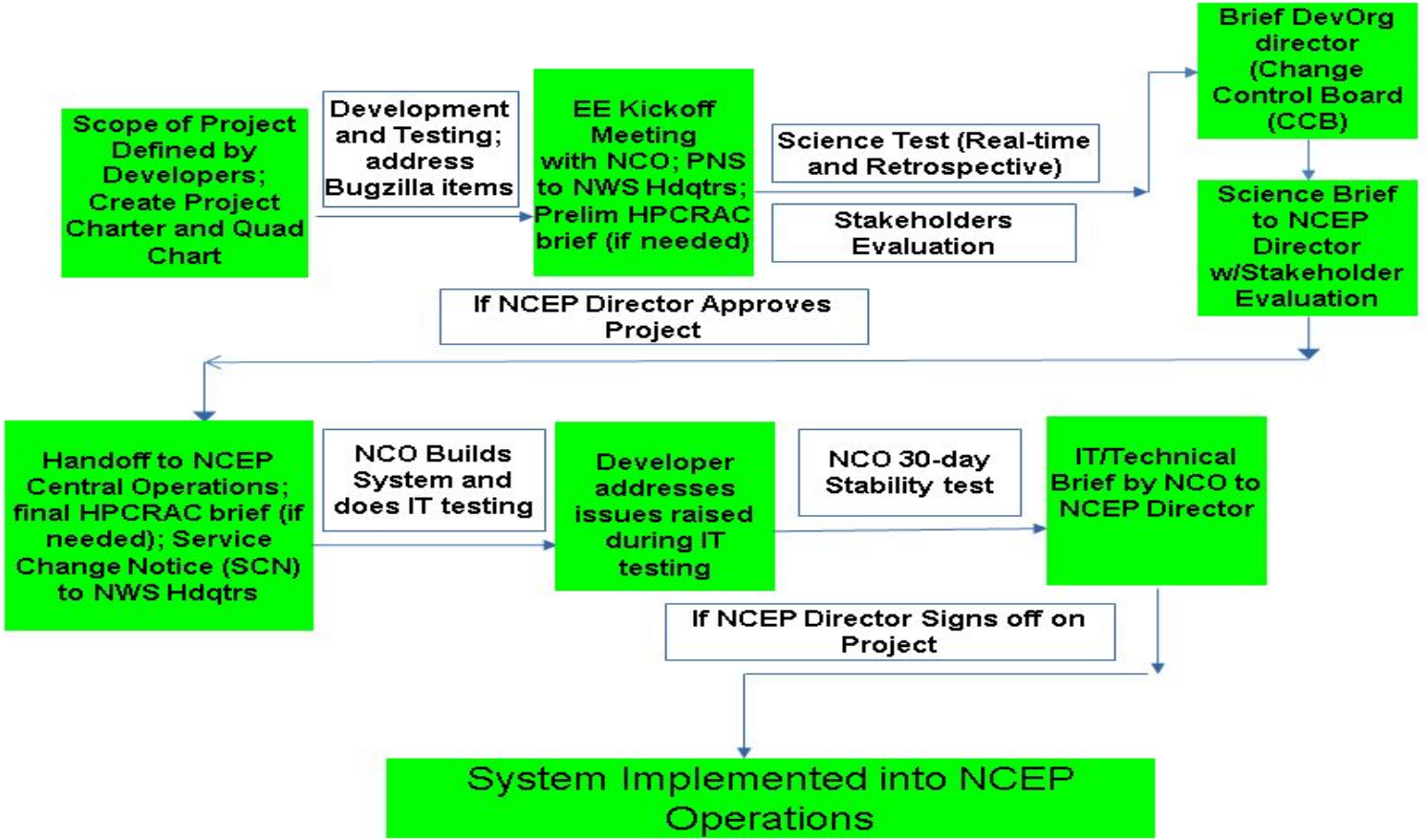
- Parallel and serial workloads
- Applications run on only 1 component of the system
- Machines have LSF (IBMs and Dells) and ALPS/LSF (Cray) as job schedulers
- Priority queues
- Run production suite using ecflow. Working to get developers using ecflow.
- Use “white space” on production for additional development work
- Upgrades have been opportunity to move applications between components of WCOSS







Transition to Operations



R2O Transition Challenges

Two major challenges for transition to operations and back to development:

- Different environments
- Production Suite Complexity

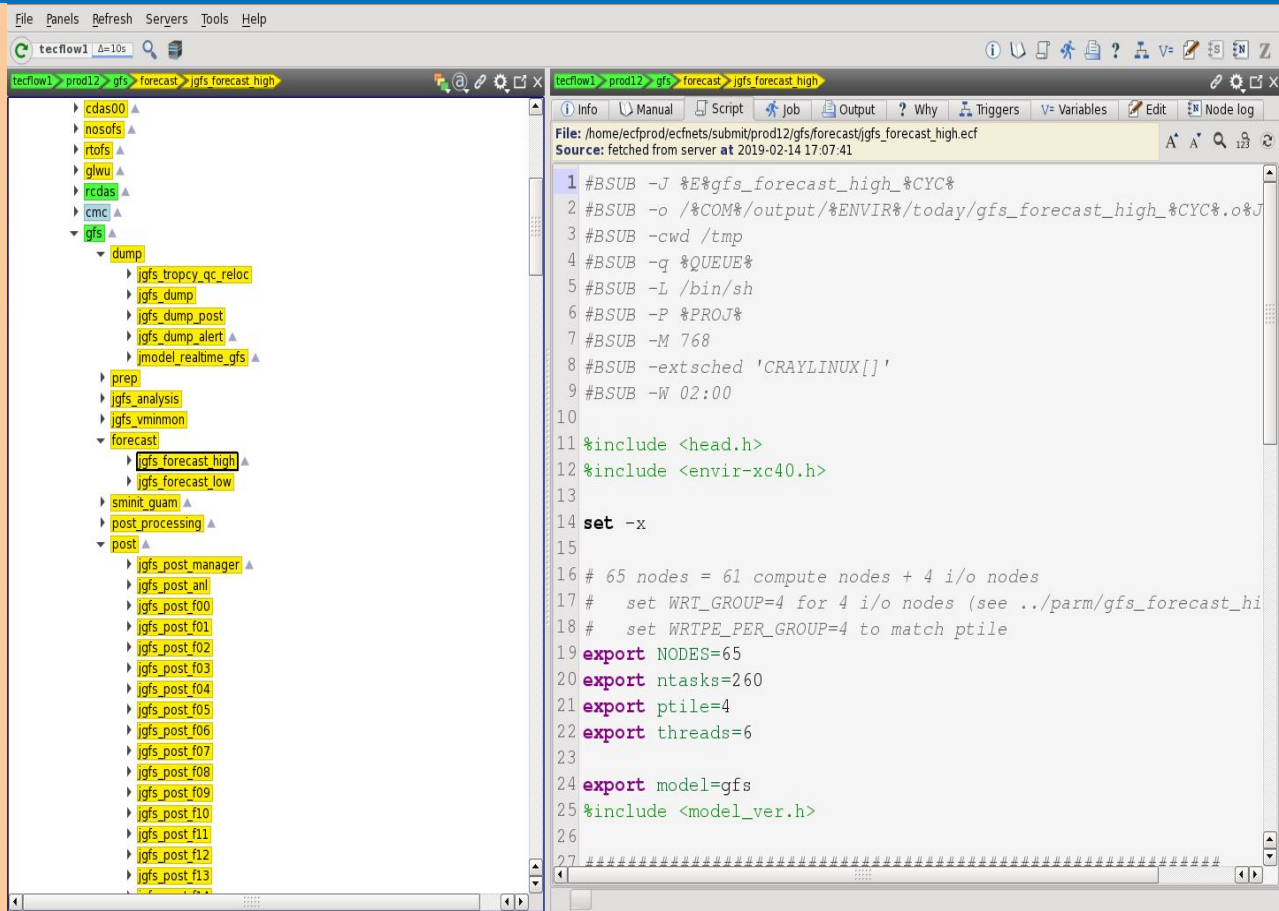
R2O Transition Challenges
Different Environments

Problem - Ops vs Dev

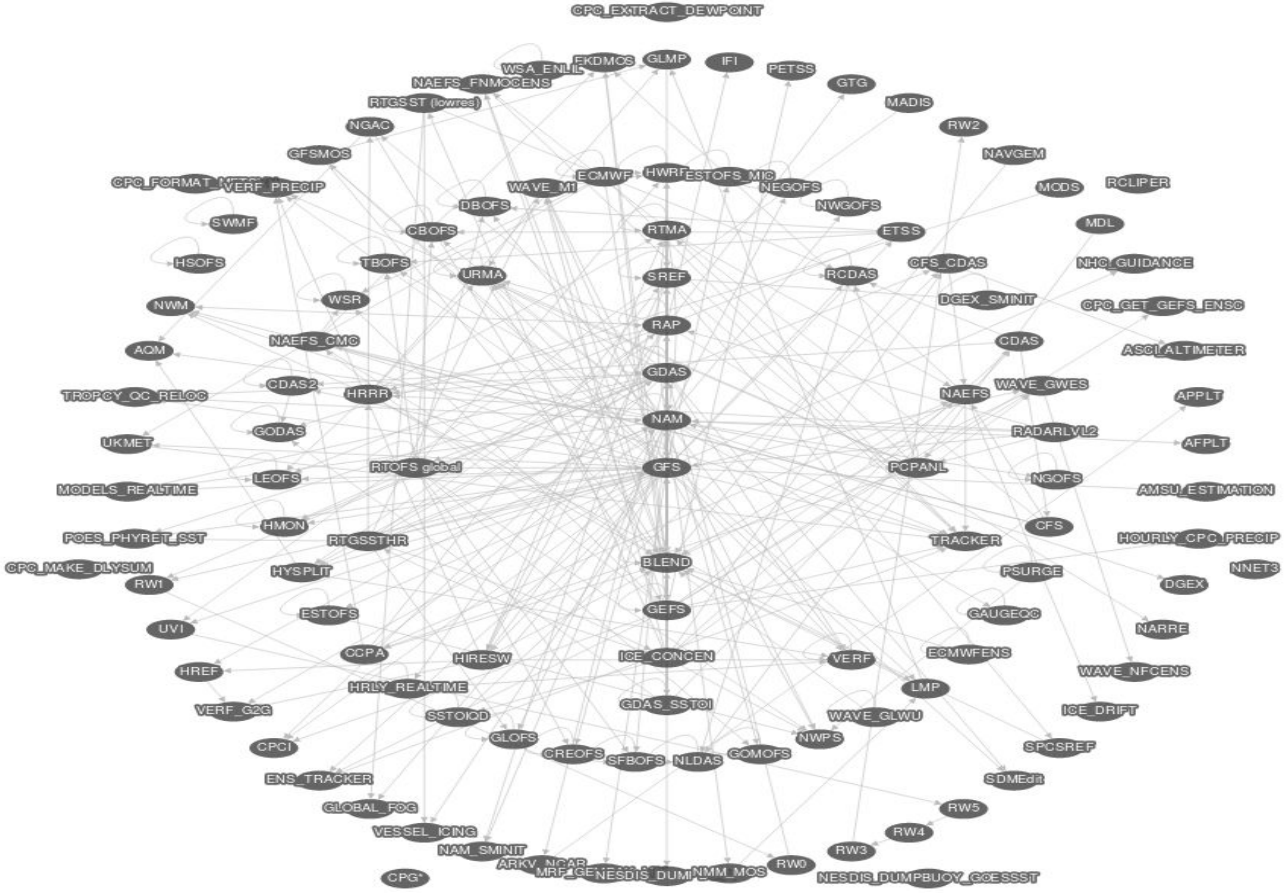
- Scheduler - cron vs ecFlow vs rocoto
- Workflow - different scripts
- System throughput / stability

What can be done?

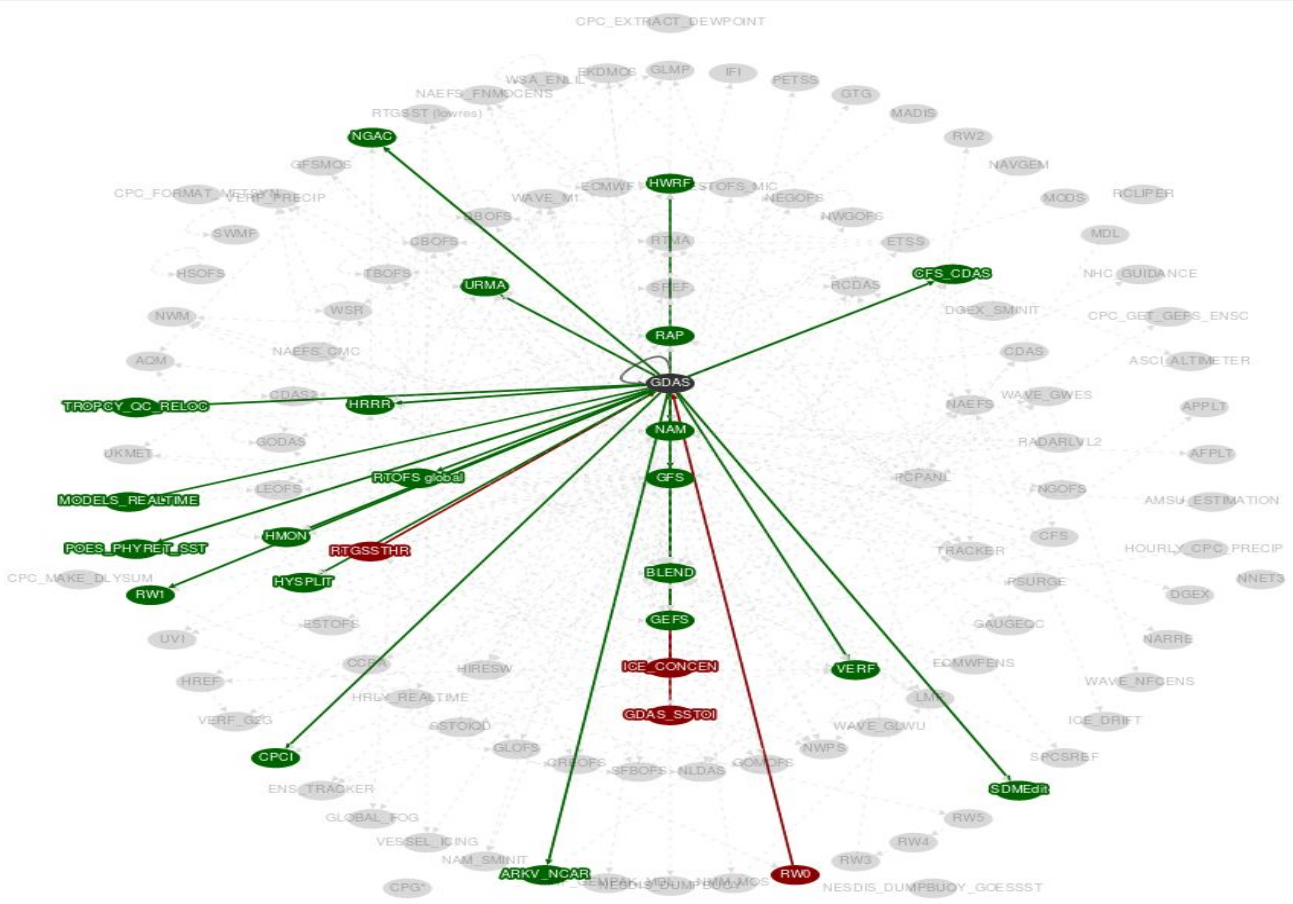
- Environment Equivalence project!
- Training underway for development ecFlow
- Development workflow will be identical to production workflow
- Fairshare scheduling in LSF



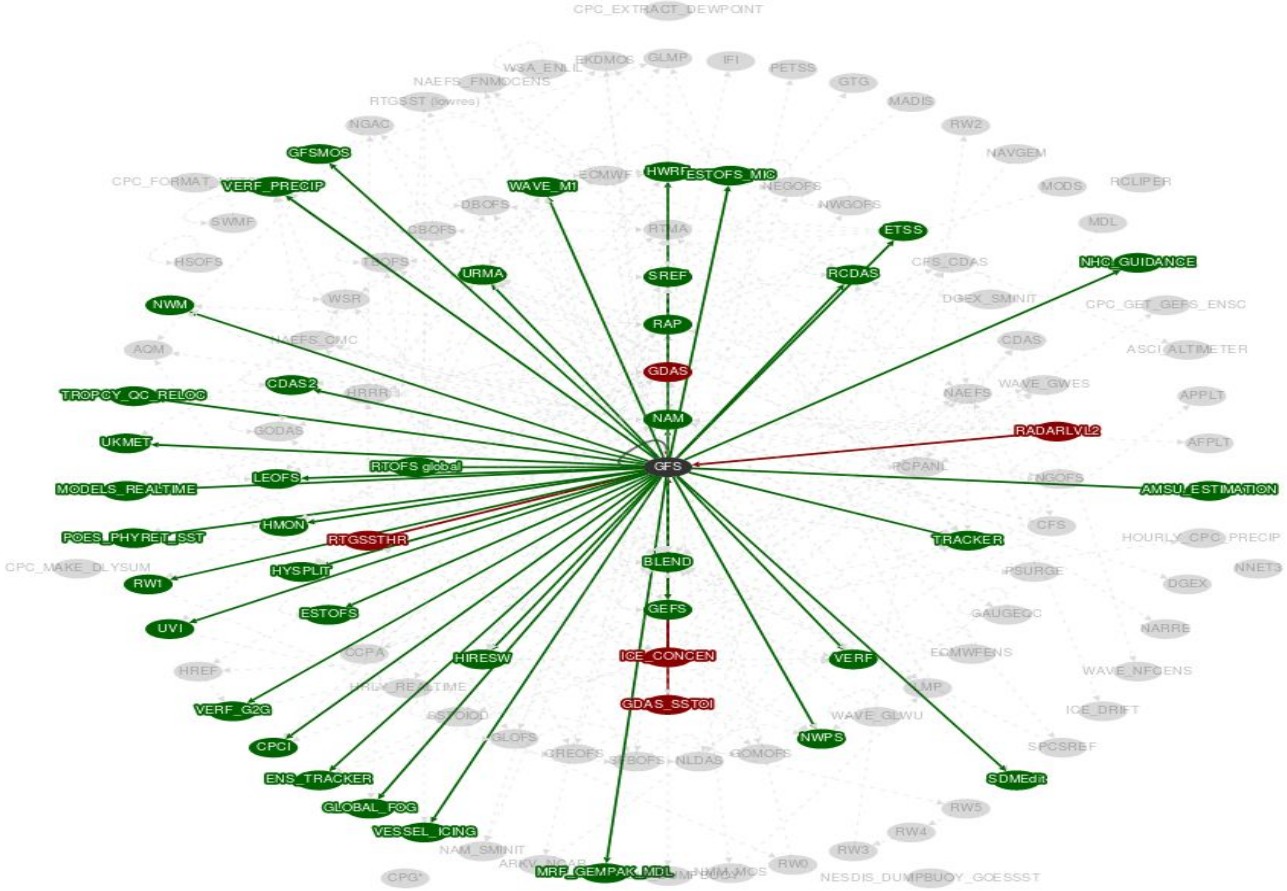
R2O and O2R Transition Challenges
Production Suite Complexity



R2O and O2R Transition Challenges
Production Suite Complexity



R2O and O2R Transition Challenges
Production Suite Complexity



R2O and O2R Transition Challenges
Production Suite Complexity

Problem - Dependencies

- Upgrade GFS → 40+ applications need to be tested and verified
- 100+ applications; most of which want yearly upgrades

What can be done?

- Production Suite Unification!
- Not just science unification - technical unification
    - Include end to end

What comes next...

- Working on our next 10-year computing contract
  - Conducted RFI and met with industry. Next up is RFP
  - What is the best computing system for our type of workload? Parallel processing versus heavy IO pre/post-processing
  - Want to foster competition
- Potential for Supplemental Funding Bills for HPC from the US Congress
- Cloud vs on-premise? Both?

Questions?

WCOSS Components	Phase I	Phase 2	CRAY	Phase 3	Totals
Compute NODES	640	1,080	2,048	1,212	4,844
Compute Racks	8.9	15	12	16.5	52
Spare Nodes	20	18	30	36	104
Peak TFs	208	572.3	2,045	1,412	4,237
Cores (Compute and Service)	9,920	25,920	50,176	36,736	122,752
Spare Cores	320	432	720	2,800	4,272
Processor Type	Intel Sandy Bridge	Intel Ivy Bridge	Haswell & Sandy Bridge	Intel Broadwell	
Processor Clock Speed	2.6 Ghz	2.7 Ghz	2.6 Ghz	2.6 Ghz	
Cores/node	16	24	24	28	
Service Nodes	54	58	100	100	312
Memory/core	2 GiB	2.66 GiB	2 GiB	2 GiB	
Disk Storage (useable PB)	2.59	2.034	3.5	5.49	14
Shared Storage (TB)	259	266			525
Backup Tape Capacity (TB)	600				600
Interconnect Fabric	Mellanox FDR	Mellanox FDR	Mellanox FDR	EDR Infiniband	
Operating System	Red Hat Linux	Red Hat Linux	CRAY and SUSE Linux	Red Hat Linux	
Filesystem	GPFS	GPFS	GPFS	GPFS	
Workflow Manager	LSF	LSF	LSF/ALPS	LSF	
Workflow Scheduler	ecFLOW	ecFLOW	ecFLOW	ecFLOW	