Supercomputing at the United States National Weather Service (NWS)

Rebecca Cosgrove

Deputy Director, NCEP Central Operations

United States National Weather Service

18th Workshop on HPC in Meteorology

September 26, 2018

Outline

- NWS/NCEP and NCEP Central Operations
- Numerical Modeling in the NWS
- Supercomputing in the NWS
- Today’s Computers -- WCOSS
- Lessons Learned
- The Path Forward

NWS National Centers for Environmental Prediction

Specialized Services – Common Mission

- 490 FTE
- 237 Contractors
- 20 visitors
- 5 NOAA Corps Officers
- $137M Budget

Aviation Weather Center
Kansas City, MO

Space Weather Prediction Center
Boulder, CO

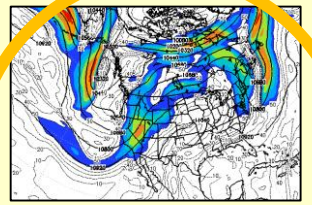
Storm Prediction Center
Norman, OK

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

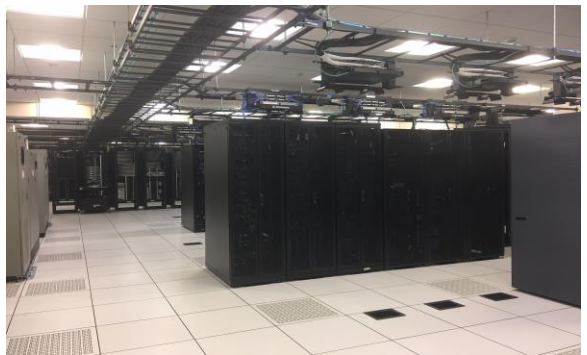
NCEP delivers national and global operational weather, water and climate products and services essential to protecting life, property and economic well-being.

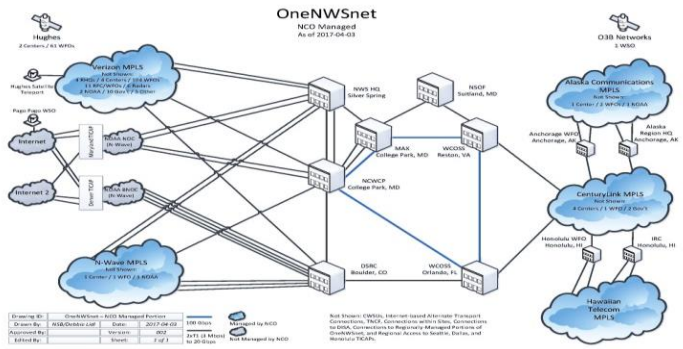
Vision

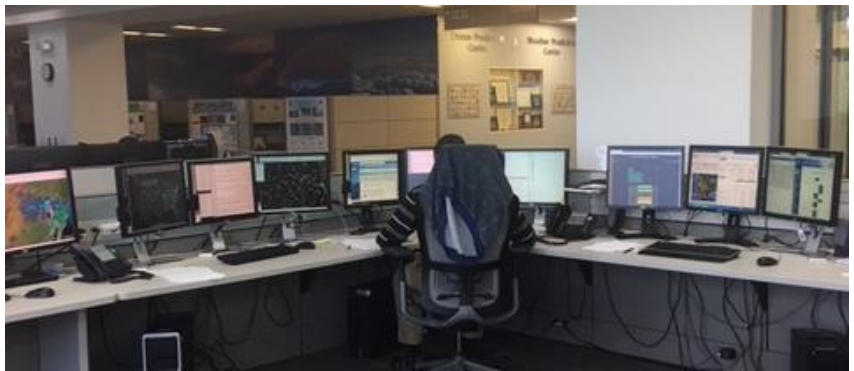
The trusted source for environmental predictions from the sun to the sea, when it matters most.

NCEP Central Operations (NCO)

- Operational areas of responsibility:
  - Maintain communications networks for NCEP Centers, NCO's 4 data centers, and all of the NWS Weather Forecast Field Offices
  - Ingest and disseminate all data for the NWS, including running the NWSTG and functioning as GISC Washington
  - Provide 24x7 Tier 1 monitoring and support for operational systems across the NWS.
  - Procure and operate the Integrated Dissemination Program (IDP) virtual infrastructure compute systems and applications running on them.
  - **Procure and operate NOAA's two operational high performance computing clusters and the NCEP Production Suite (NPS) running on them**

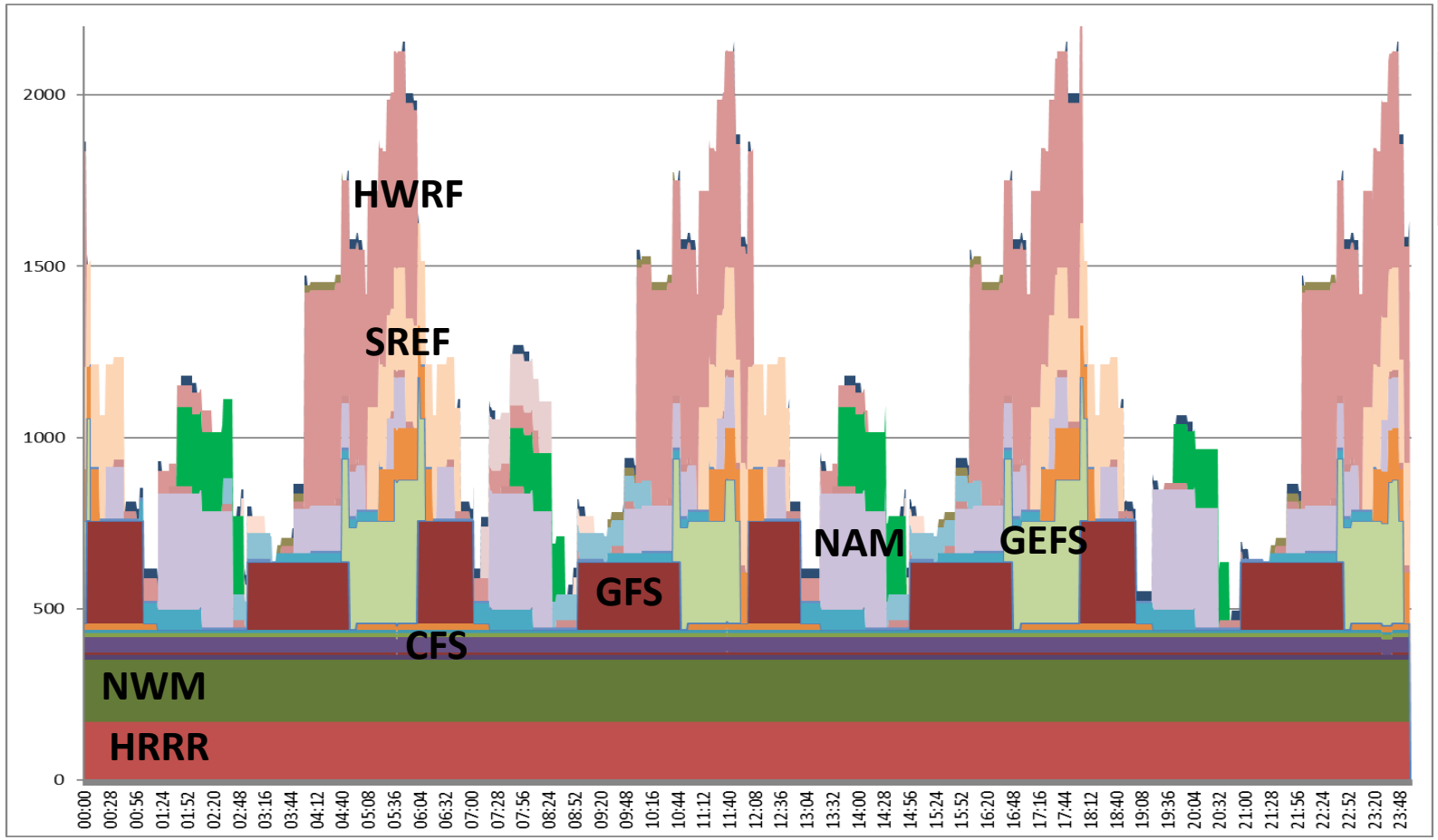


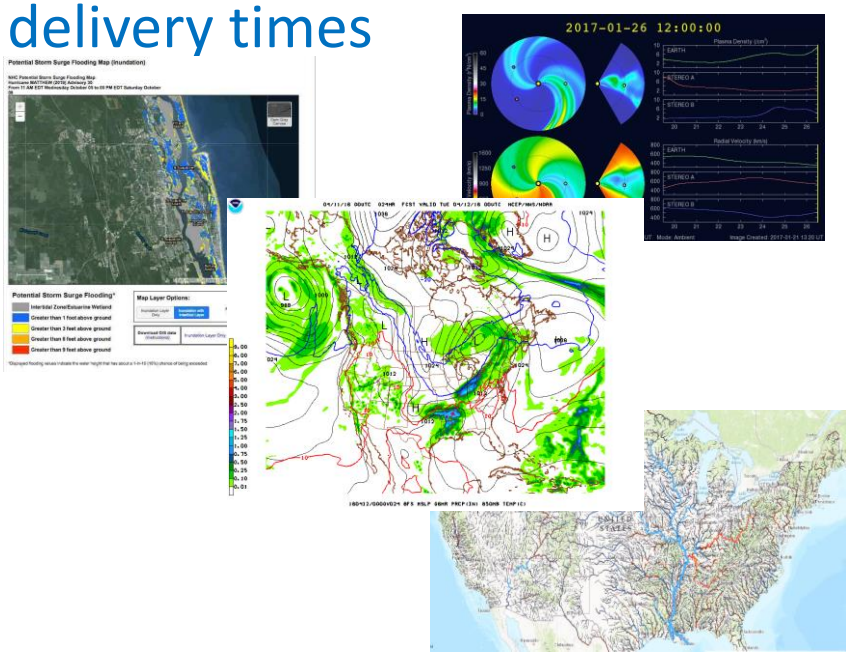




NCEP Production Suite (NPS)

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

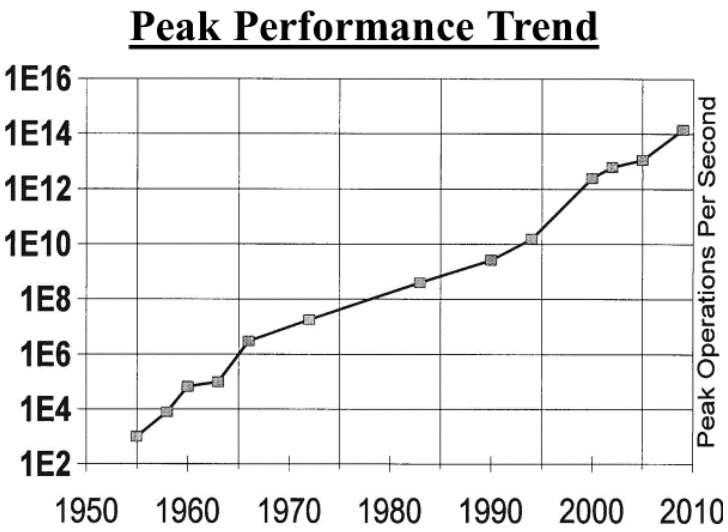


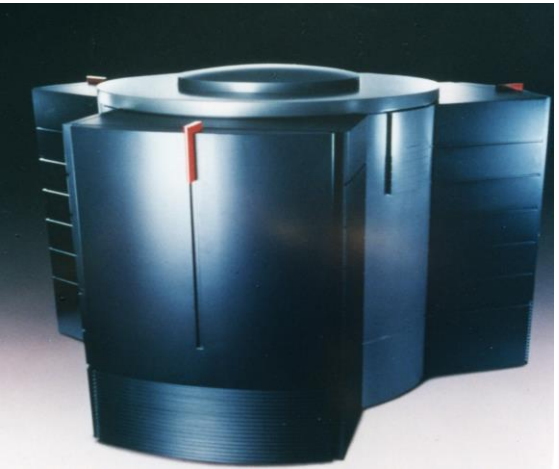


GFS – Global Forecast System
CFS – Climate Forecast System
NWM – National Water Model
HRRR – High Resolution Rapid Refresh
GEFS – Global Ensemble Forecast System
NAM – North American Model
SREF – Short-Range Ensemble Forecast
HWRF – Hurricane Model

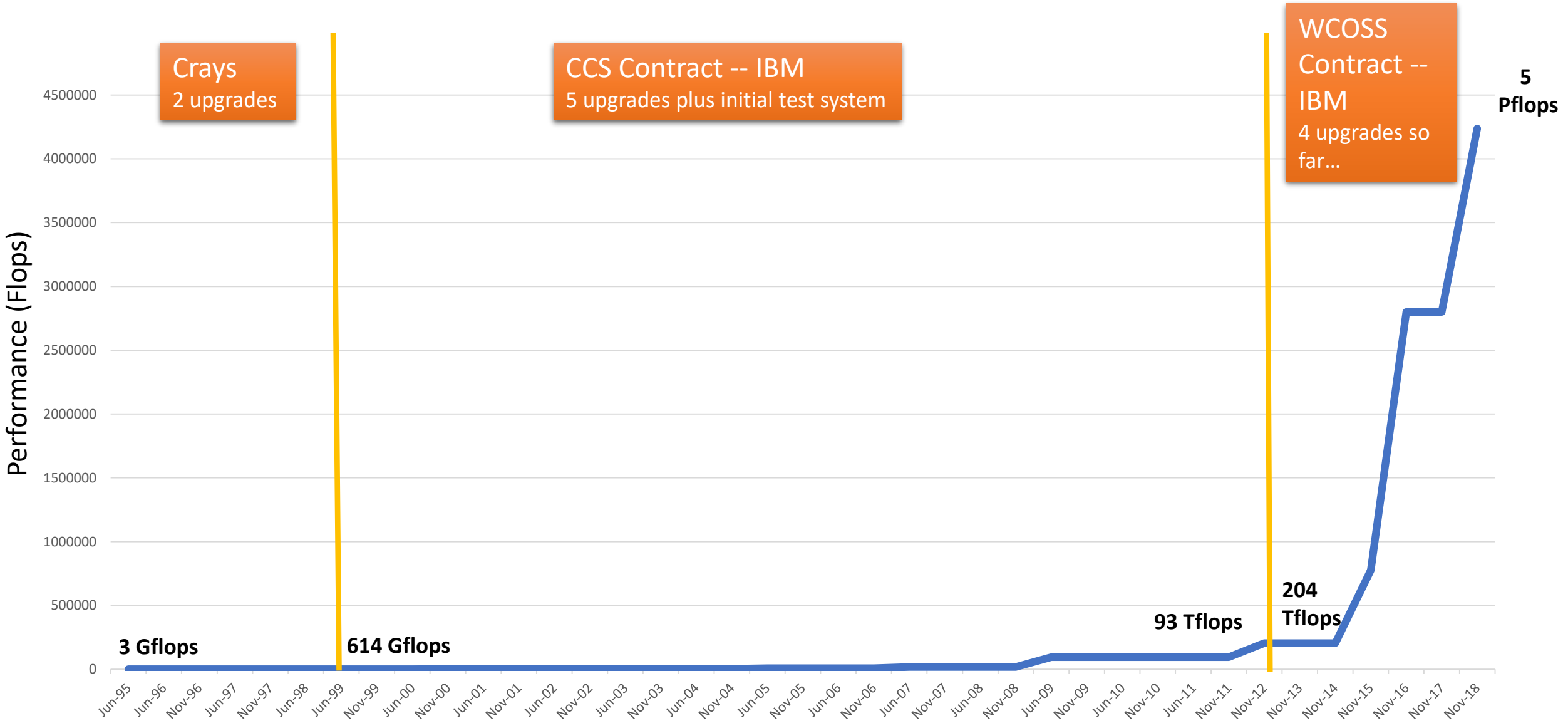
Supercomputing in the NWS

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



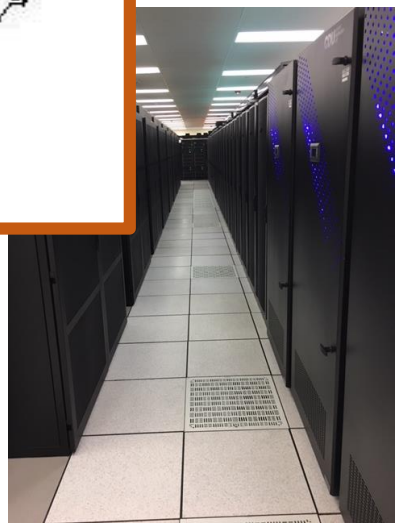


NWS Operational Computing Over the Last 20 Years



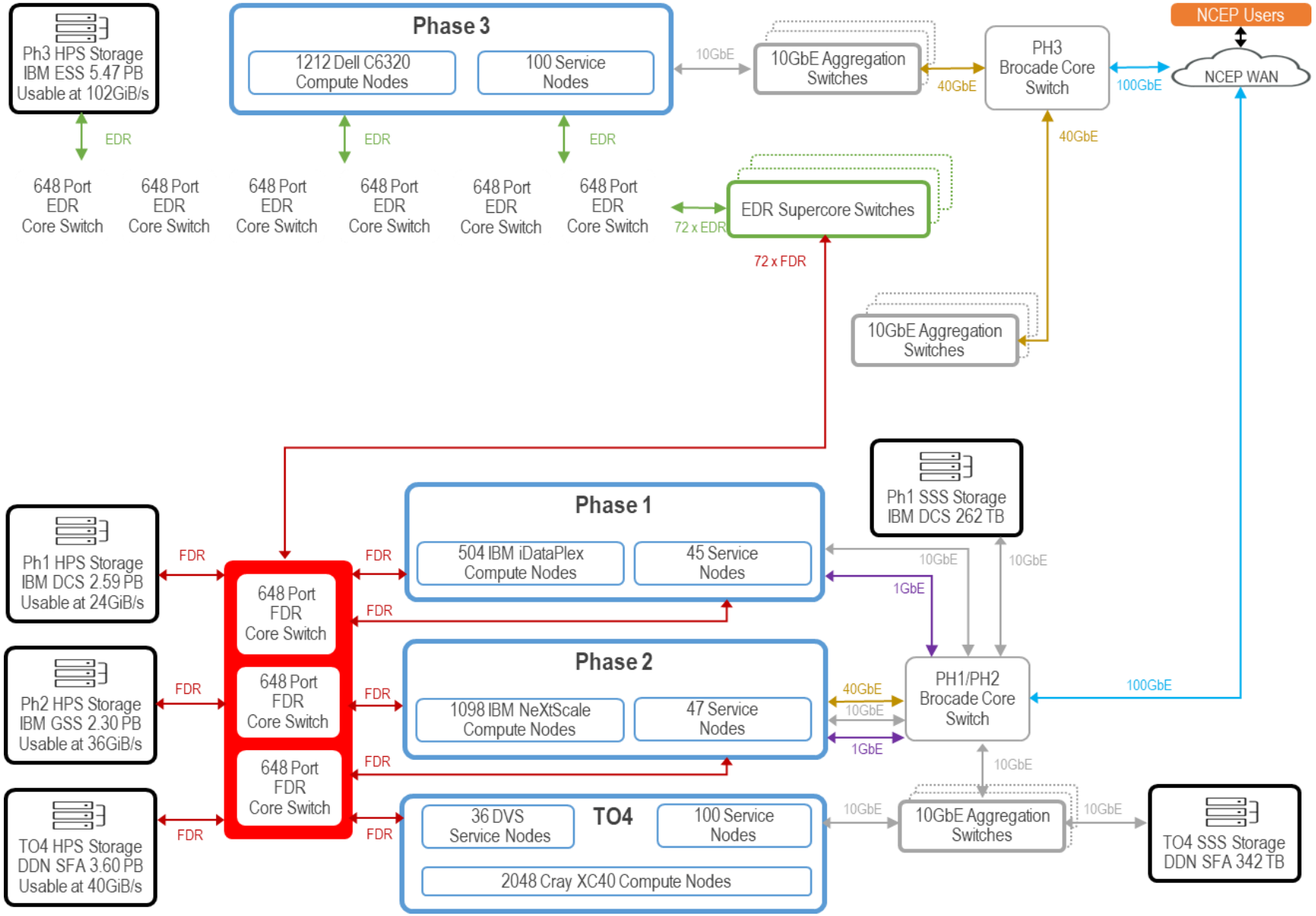
Current Computing -- WCOSS
Weather and Climate Operational Supercomputing System

- 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 4.98 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
- Processes 3.5 billion obs/day, produces 140 million products/day, distributes over 10 TB of guidance/day
- FISMA (Federal Information Security Management Act) High System

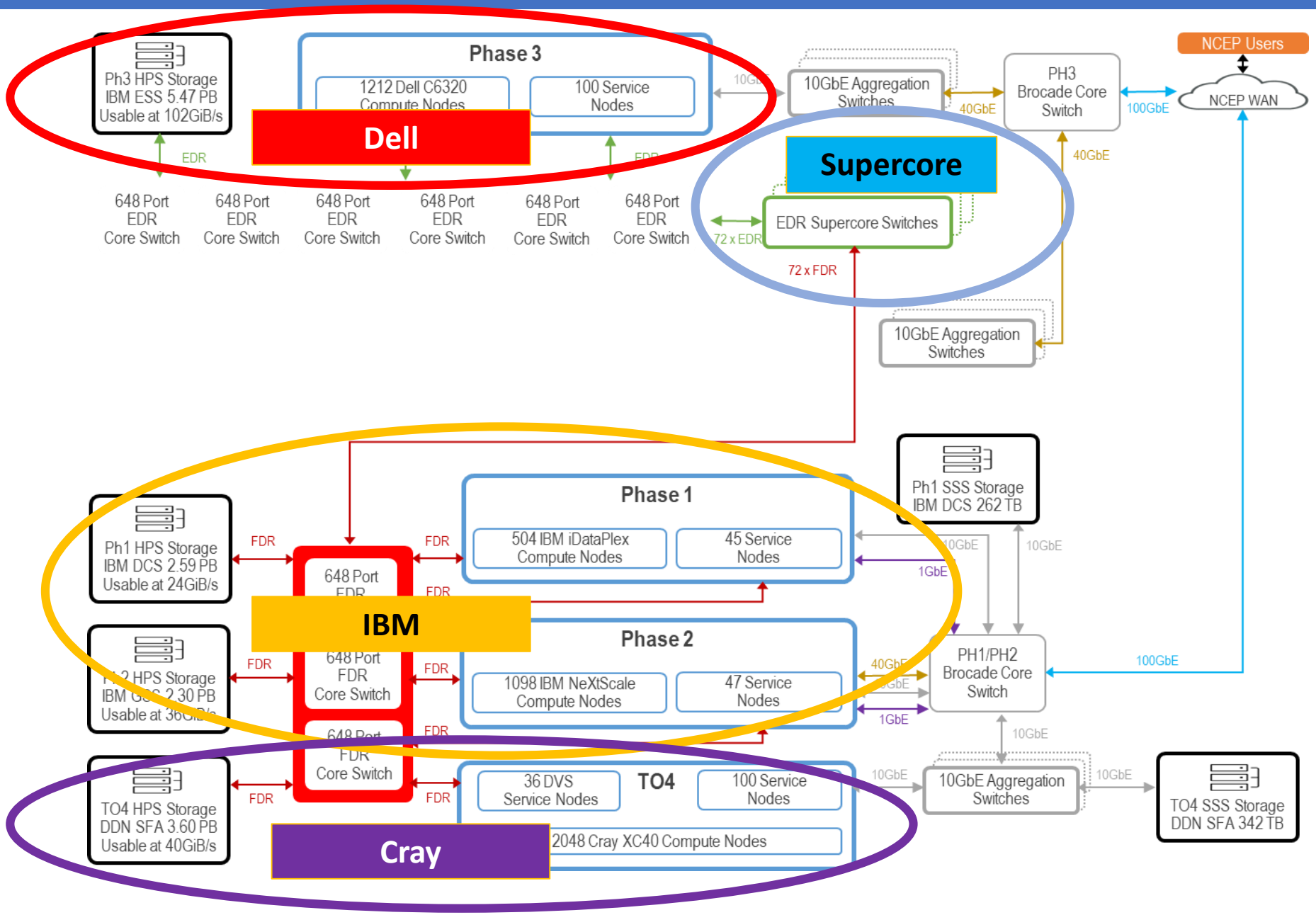


WCOSS Schematic

- 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

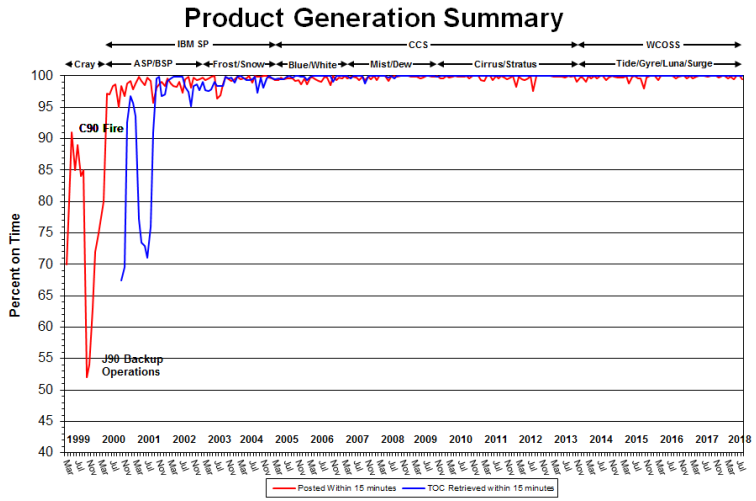


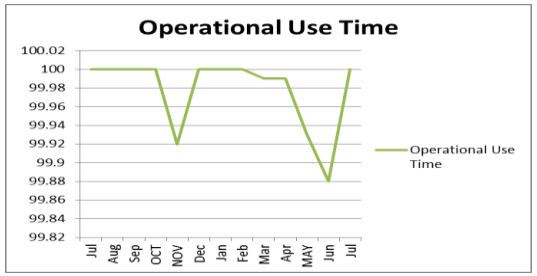
WCOSS -- Providing Operational Resiliency/Reliability

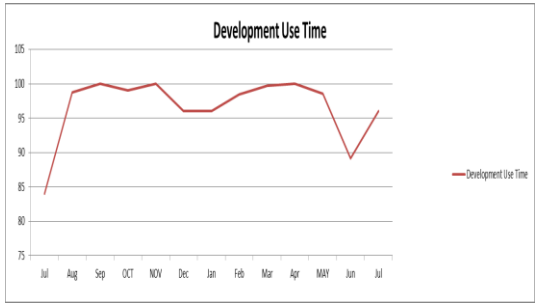
Performance Requirements

- 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







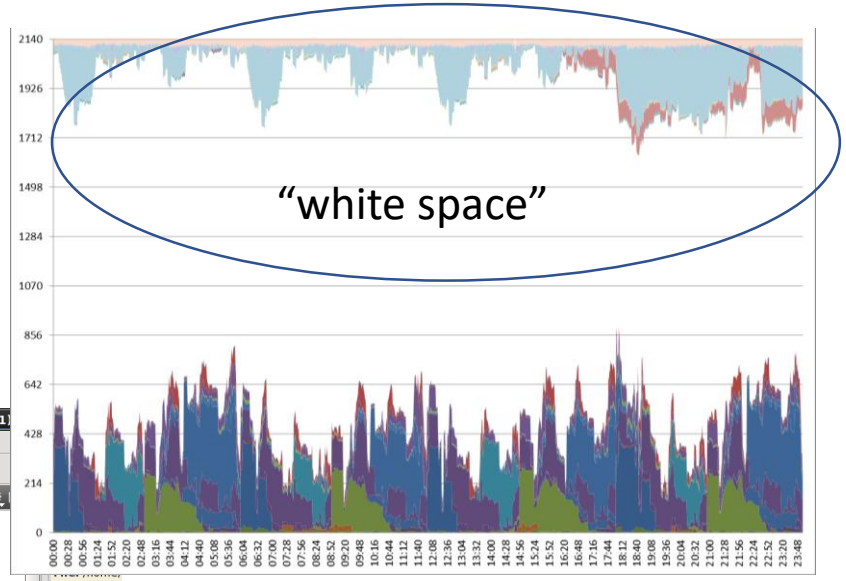
Two Identical Systems – one Production and one Development

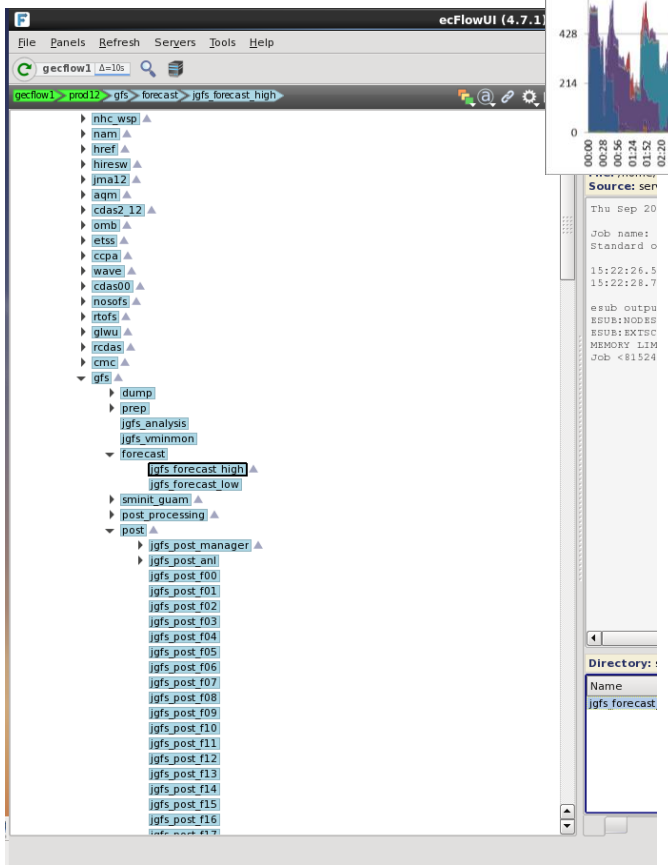
- Production locked down to NCO and select users
- Development machine is open to all users
- Switch between systems takes 15 minutes, but models bleed off for a few hours.
- Data mirrored between two systems – 40TB per day. Must have that for failovers and development work.
- For major maintenance activities,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

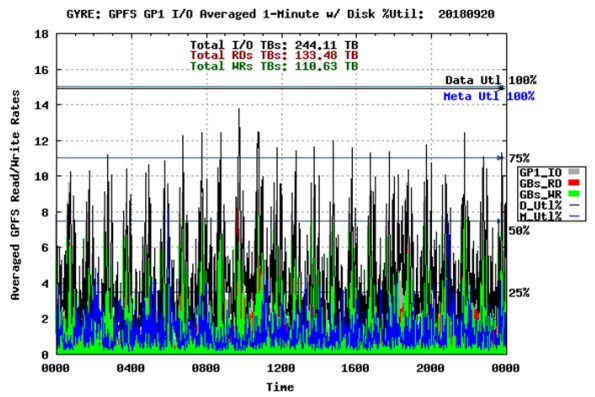
WCOSS – Workload Management

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



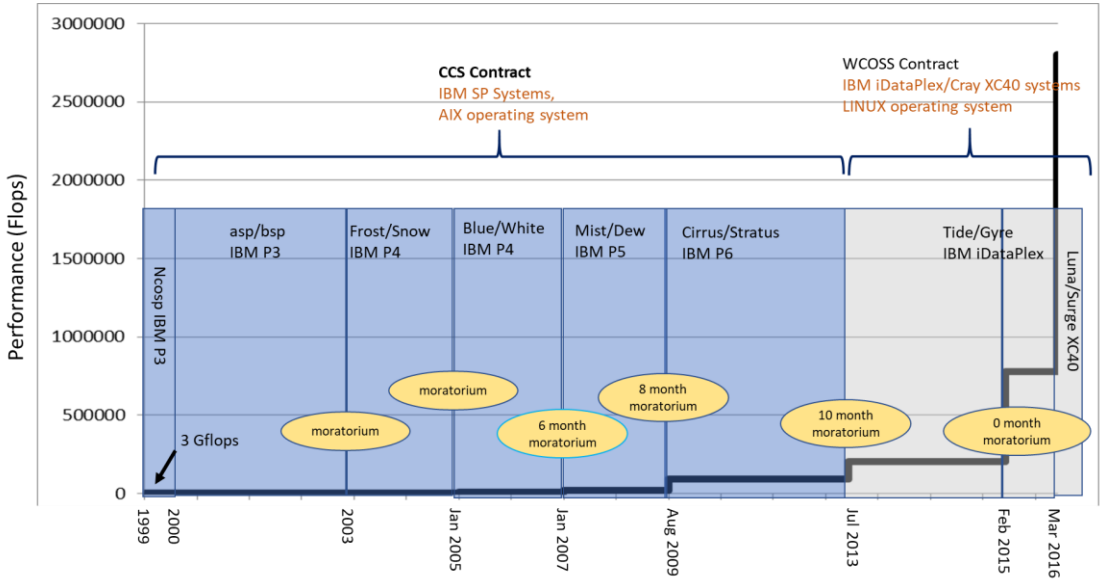




WCOSS Transitions

- Previous contract had forklift replacements over the 10-year contract. WCOSS has had add-ons which led to heterogeneous systems.
- Pros and cons to both approaches

Forklift	Issue	Add-on
☹	Model upgrades/development during transition	☺
☺	All efforts focused on the transition	☹
☺	Resulting system can have different architectures	☹ \ ☺
☹ \ ☺	All codes must be ported. But also allows for removing legacy codes	☹ \ ☺
☺	Work on installing/testing new system can impact operational system including outages to grow (or later remove) phases of the system	☹
☺	Entire system has same architecture and models can be run on entire system	☹
☺	Facilities – already there probably for forklift, but is it big enough for the add-on for the life of the contract?	☹
☹	Overlap of two systems requires more power	☺
☹	Can respond relatively quickly to unexpected funding, like Sandy Supplemental	☺



Lessons Learned From WCOSS

- **Technical Challenges**
  - IO is our bottleneck now – codes not optimized for them
  - Shared storage – has never met our needs to keep data synced between the two systems.
  - Took 2 years to enable caching on the Crays – applications had variable runtime and delayed product delivery
  - Never should have bought Crays with GPFS
- **Physical Issues**
  - Do you have enough space? Can you grow?
  - Will any vendors have requirements for how systems are laid out on the floor?
  - Having robust data centers makes a difference – good for gov't, can be impediment to bids
  - Cabling issues
- **Contract Matters**
  - Contract limitation – fixed price contract, but with a requirement for growth of compute.
  - Penalties in the contract – having defined penalties and formulas has worked well
  - Spare nodes – have encouraged concurrent maintenance
  - Good metrics/benchmarks are critical. The more we describe the work we do, the better machine we should get.
  - Very important to have customer-focused PM and top notch lead system administrator

Lessons Learned From WCOSS

- **Operations and Maintenance**
  - System repairs – do you have the equipment on hand or local to conduct repairs?
  - System metrics are critical for optimizing use of these large systems
  - Critical to maintain consistent configurations between two systems
- **System Configuration**
  - Separate your data and metadata
  - Filesystem Structure – fastest recovery/maximize IO/split production and development workloads to provide stability?
  - Acceptance – don't run through acceptance in a configuration that you won't operate with
- **Heterogeneous**
  - Can end up with split support team -- prime and sub-contractor communication and process issues.
  - Heterogeneous systems can lead to more portable code
  - LINPAC Benchmark can't combine heterogeneous systems into one ranking
  - Is the system designed for growth?
  - Sale of X86 product line to another company mid-contract created issues
  - Original Phase 1 and Phase 2 IBM systems got too old – can you really replace them?
  - Supercore has worked well. Crossing components hasn't overwhelmed systems
  - System upgrades are hard – keeping components at same/similar levels

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 tie increase in compute to industry index. Our budget is fixed, but compute costs change over the life of a contract
  - Want to foster competition
- Potential for Supplemental Funding Bills for HPC from the US Congress
- What do we do with our Cray?
  - The IBM and Dells are leased, Cray is not.
- Operations is 24x7x365

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	

Rebecca.Cosgrove@noaa.gov