



# HAFS Computations

**Sundararaman Gopalakrishnan**  
**Frank Indiviglio**  
**Vijay Tallapragada**  
**Ghassan J. Alaka, Jr.**

2020 HFIP Annual Meeting  
19 November 2020

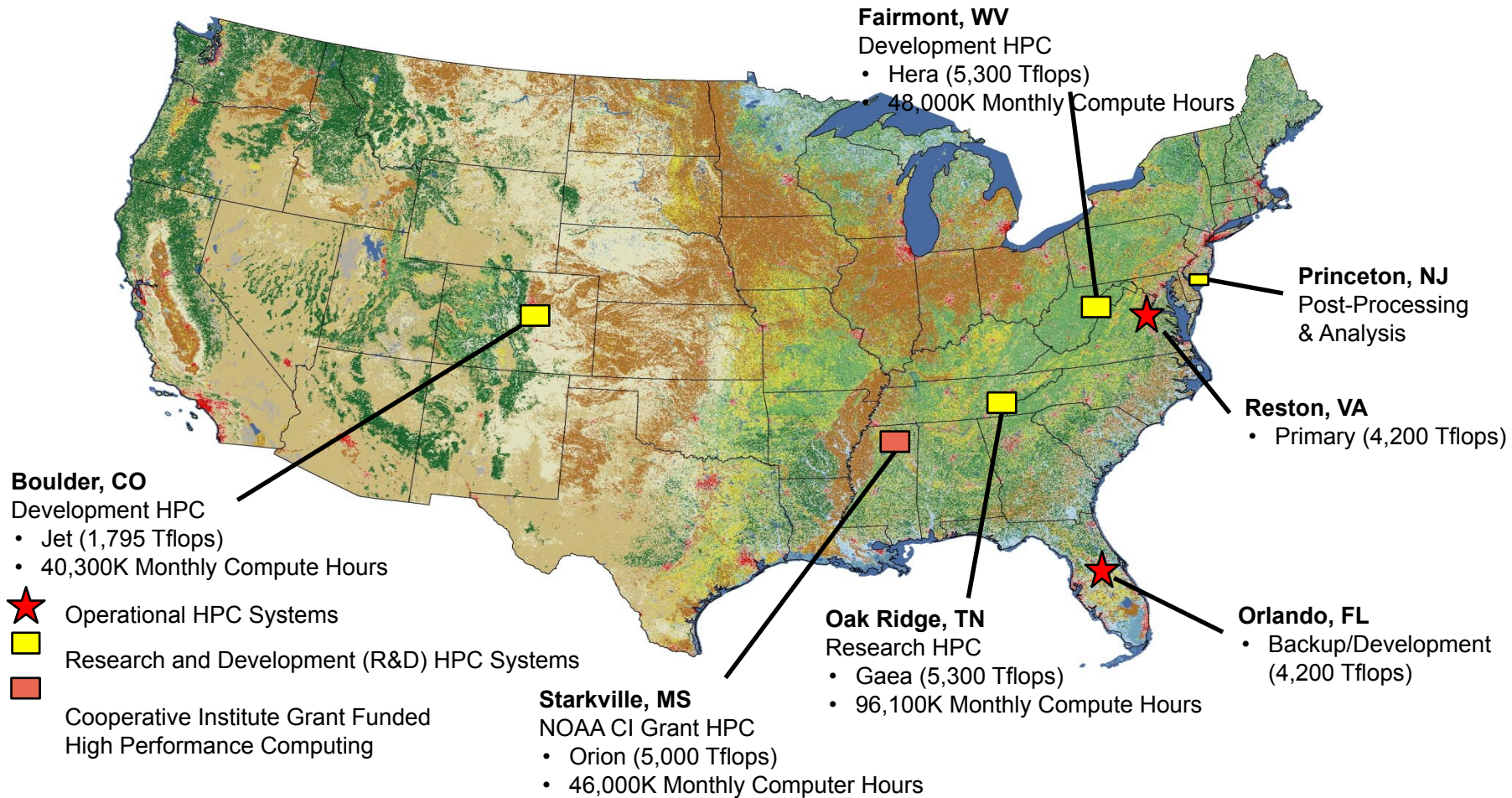


# NOAA RDHPCS

Frank Indiviglio

Deputy Director High Performance Computing and Communications, OCIO

# NOAA's High Performance Computing



# HPC Priorities

- User Priorities

- User engagement throughout implementation and delivery
  - Integrated planning input throughout (HUG, Allocation Committee, Expanded User Requirements Discussions)
  - Move to continuous integration of requirements and delivery

- Mission Focus

- Ensure systems and software are aligned to mission goals

- Modernize with technology

- Explore new technologies in hardware and software

- Make small bets for future success

- Make targeted investments in methods and technologies to improve fulfillment of the mission and modernization efforts



# NOAA RDHPCS

Vijay Tallapragada

# Reservation and workflow

- Application projects review

- 7 projects selected. Total **12,063,000** Cr-Hrs/month
- 3000 Cr-Hrs/month on sjet, 1,270,000, and the rest are on xjet
- The user workflow review started on June 15
- There are 13 CM tickets for creating/adjusting reservations

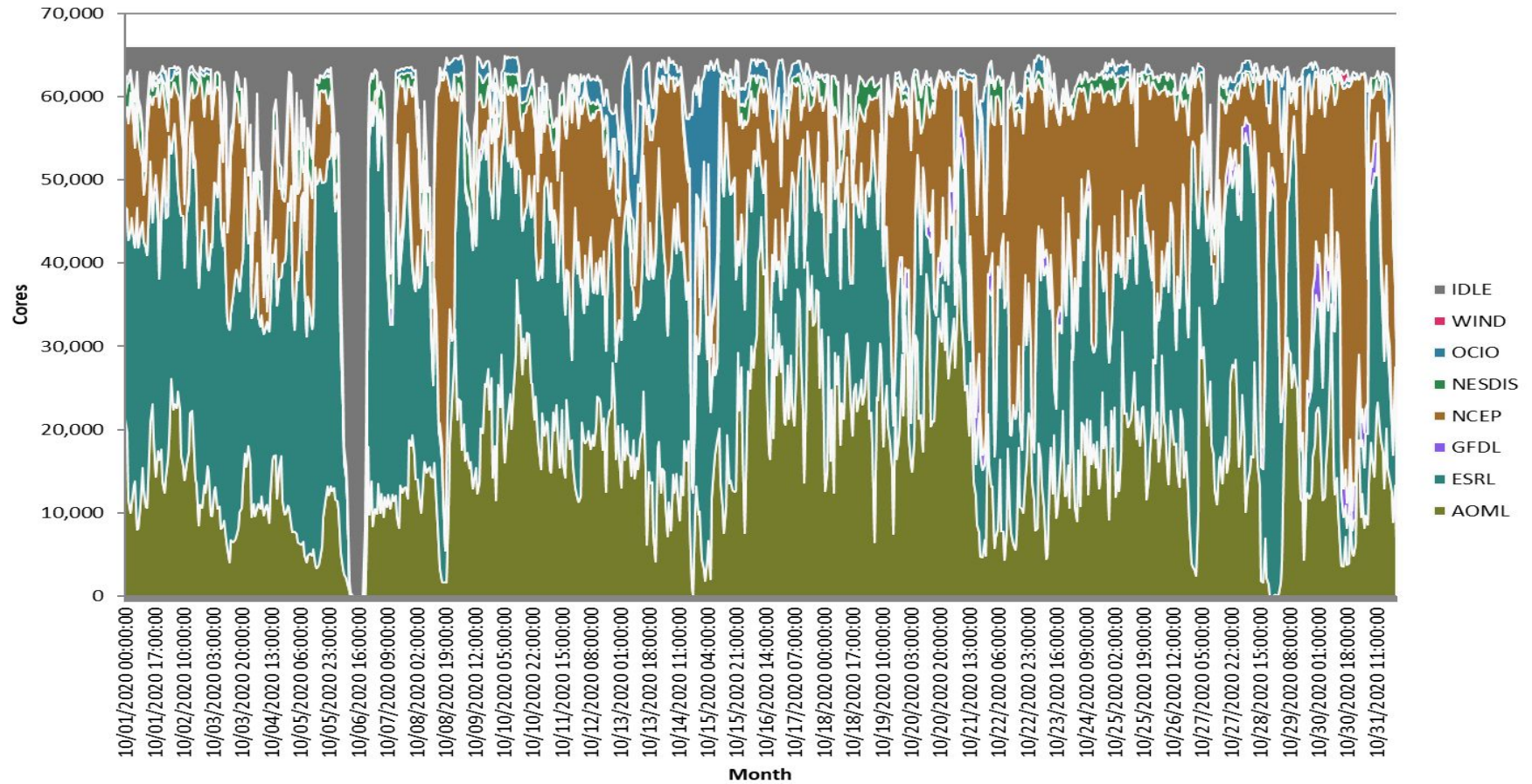
- Reservations

- Pre-production Reservation Deployment on July 03
- There are 85 reservations for HFIP projects
- All reservations are monitored every 6 hours (missing or not)
- Reservation usages have been monitored for each epoch and reviewed every week

- User tickets

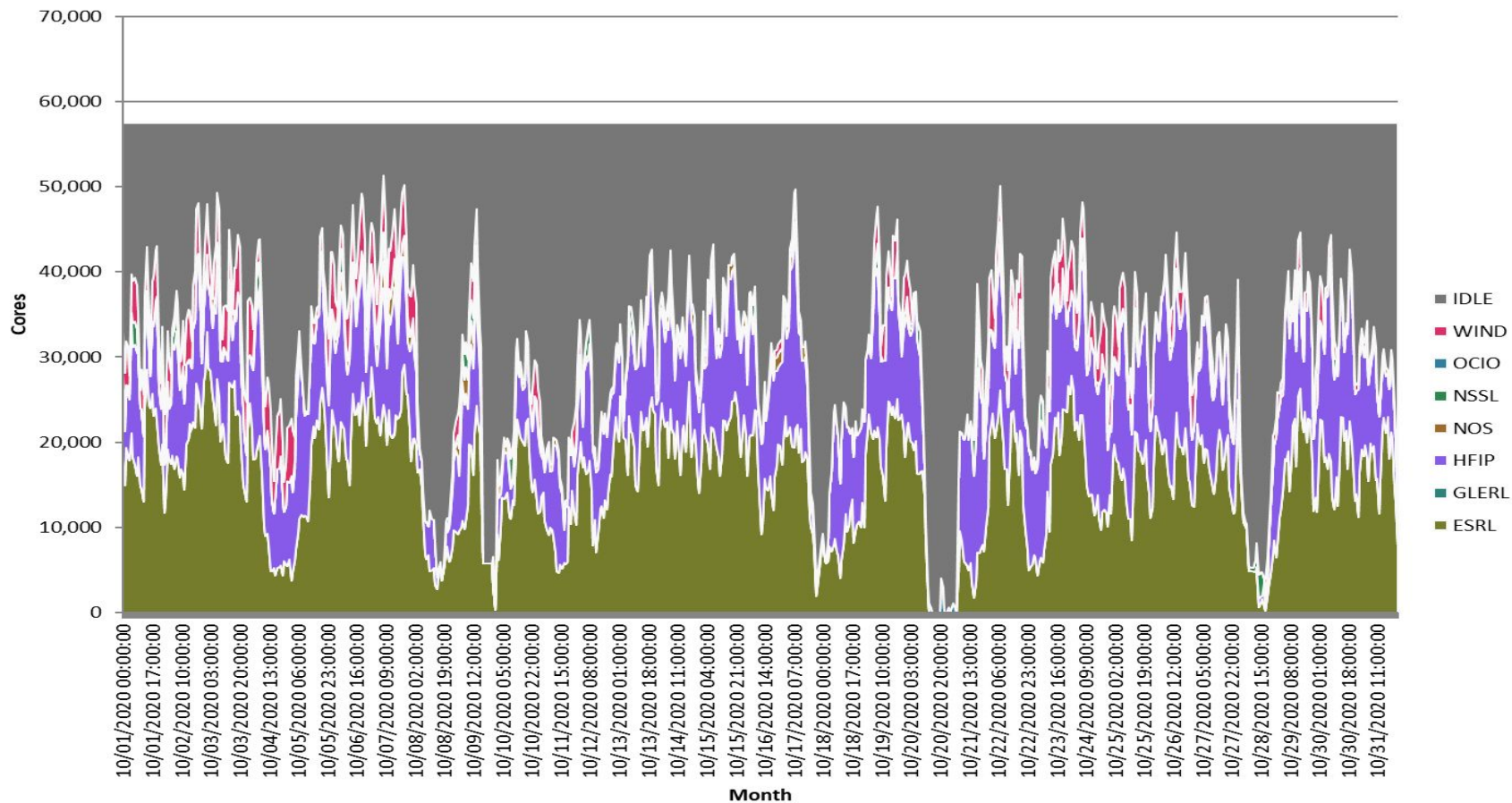
- There are 35 help tickets from HFIP reservations.

# Hera Center Usage





# Jet Center Usage





# RDHPCS: Orion

**Manufacturer:** Dell EMC

**Model:** PowerEdge C6420

**Interconnect:** Mellanox Infiniband  
HDR-100

**Processor:** Xeon Gold 6148 20C 2.4GHz

**Total Memory:** 338,688 GB

**Total Nodes:** 1,800 (1,792 Compute and  
8 Bigmem)

**Total Cores:** 72,000

**Cores per Node:** 40



**5 petaFLOPS per second, making it the  
60th most powerful supercomputer in  
the world according to Top500.org**

# The New WCOS2 Coming Online Soon

## Locations

- Manassas, VA
- Phoenix, AZ

## Performance Requirements

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

## Configuration

- Cray Shasta system
- 12.1 PetaFlops
- Multi-tiered storage
  - 2 flash filesystems each with...
    - 614 TB usable storage
    - 300 GB/s bandwidth
  - 2 HDD filesystems each with...
    - 12.5 PB usable storage
    - 200 GB/s bandwidth
  - Total aggregate - 26.2PB at 1TB/s
- Lustre parallel filesystem
- PBSpro workload manager

- Compute nodes
  - 2,560 nodes (60 spare)
  - 327,680 cores
    - 128 cores/node
  - 1.3 PB of memory
    - 512 GB/node
- Pre/post-processing nodes
  - 132 nodes (4 spare)
  - 8,448 cores
    - 64 cores/node
  - 132 TB of memory
    - 1TB/node
- 200Gb/s Slingshot interconnect

# FY18 HSUP Allocations

Portfolio	Project	RDHPCS PI	Hera/Jet Allocation (M)	Orion Allocation	Total
BMC					
	fy18-dras-1A1	Curtis Alexander	0.5	0.5	1
	fy18-dras-infra	Arun Chawla	0	0.2	0.2
	fy18-dras-ens	Jeff Whitaker	6	2	8
TOTAL			6.5	2.7	9.2
AOML					
	fy18-dras-1A4	Gopal S.	0	10	10
	fy18-dras-3A1	Frank Marks (Andrew Hazelton)	0	5	5
	fy18-dras-4A2	Lidia Cucurull	3.75	1.25	5
TOTAL			3.75	16.25	20
NCEPDEV					
	fy18-dras-physics	Vijay Tallapragada	0	4	4
	fy18-dras-hafs	Avichal Mehra	2.5	7.5	10
	fy18-dras-efso	Daryl Kleist	2	0	2
	fy18-dras-jcsda	Daryl Kleist	2.5	7.5	10
	fy18-dras-obsproc	Daryl Kleist	0.2	0	0.2
TOTAL			7.2	19	26.2
COASTAL					
	fy18-dras-3A3	Nicole Kurkowski	0	0.1	0.1
TOTAL			0	0.1	0.1
TOTALS			17.45	38.05	55.5

# FY18+FY19 HSUP Allocations

Portfolio	Project				
		A			B
		FY18 ALLOCATIONS: Core-hours per month (M)	FY19 REQUIREMENTS: Core-hours per month (M)	FY18 ALLOCATIONS + FY19 REQUIREMENTS project Core-hours per month (M)	RECOMMENDED ALLOCATION: (FY18+FY19)*.82 Core-hours per month (M)
HPC Portfolio	Projects				
	fy18-dras-3A3				
TOTAL		0.10	0.00	0.10	0.08
NESDIS					
	fy19-dras-aida-HU-3				
TOTAL		0.00	2.00	2.00	1.64
NWSOWP					
	fy19-dras-FL-2				
TOTAL		0.00	0.02	0.02	0.02
TOTALS		55.50	24.52	80.02	65.50



# User Perspectives

Ghassan J. Alaka, Jr., Andrew Hazelton, Lew Gramer, Levi Cowan



# HPC User Issues

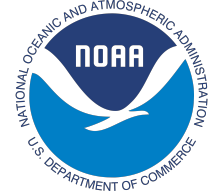
- Jet had a lot of problems this year → I/O, nodes, LFS, other hardware, etc.
  - ~10% of HAFS-B forecasts outside of the Jet reservation
  - ~35% of HWRF-B forecasts outside of the Jet reservation
  - Reservation was unreliable for several days, including gaps of 7 and 14(!) days for HWRF-B
  - Old partitions (tjet/ujet/sjet) are difficult (impossible?) to use for current models
- HPC inconsistency was very challenging, especially for a cycled models
  - HWRF-B forecast delivery was slowed down significantly because it can't skip cycles
  - HAFS forecasts that failed on Jet were completed on Hera, but this will not be easy when DA is added to HAFS
- Pls spent too much time “babysitting” real-time forecasts
  - Up to 50% of work time to keep these experiments moving along
  - Admirable effort, but unsustainable
  - Releases in Reservation system not “user friendly” - easy to give up future reservations
- Low reliability led to late delivery of ATCFs and products for important cases



# HPC User Issues

- Lack of disk space on HERA
- Fair Share policy as currently designed is "unfair". It penalizes those users who manage to use their allocated resources every month.
- HPSS access on ORION
- Access to dump data on ORION
- stmp space gets filled up quickly
- Pls spent too much time “babysitting” real-time forecasts
  - Up to 50% of work time to keep these experiments moving along
  - Admirable effort, but unsustainable
  - Releases in Reservation system not “user friendly” - easy to give up future reservations
- Low reliability led to late delivery of ATCFs and products for important cases





# Ideas for Improvement

- Reservations on Orion or Hera
  - Both are newer machines than Jet
- More resources for fallback options
  - If “HPC1” is unavailable, the forecast can run on “HPC2” with a delay
  - Reproducibility across multiple machines is important
- Full data access for DA on Orion (in progress)
  - Ability to test input data delivery and timing before new experiments start is vital
- Continue to build the relationship with RDHPCS
  - We have taken some big steps in the last few years, including improved communication
  - Fewer system changes during the season (modules, libraries, aliases, etc.)
  - Extraordinary efforts to resolve problems as quickly as possible



# Ideas for Improvement

- Fair Share policy should encourage users to meet their allocations every month. Change the calculation of Fair Share Index
- Access to HPSS is needed for archiving and fetching data
- Implement a stmp disk scrubber
- Disproportional strengths of operational computing (~25PF) and RDHPCS (15PF including Gaea)
- HFIP's success would have been impossible without Jet - refresh it to sustain the community effort!
- Cloud Computing (on demand, surge)???
- What happens after HSUP? Can HFIP retain those allocations on Hera and Orion?