U.S. NAVAL RESEARCH LABORATORY

Modernizing U.S. Navy NWP Operations: Toward Distributed HPC

Timothy Whitcomb
NRL Marine Meteorology Division
20180926T1230

U.S. NAVAL RESEARCH LABORATORY

Modernizing U.S. Navy NWP Operations: Toward Distributed HPC

Timothy Whitcomb(1), Daniel Arevalo(2)
(1) NRL Marine Meteorology Division, (2) DeVine Consulting
20180926T1230

Toward Future Navy Modeling

Navy METOC mission, basic and applied research, National ESPC, future coordinated national modeling

2020s

Bridge to NEPTUNE

- Further development, evaluation, and integration of NUOPC/ESMF-compliant community physical parameterization suites and additional coupled components.
- Emphasize performance for next-generation compute platforms
- Committed to partnerships

2019

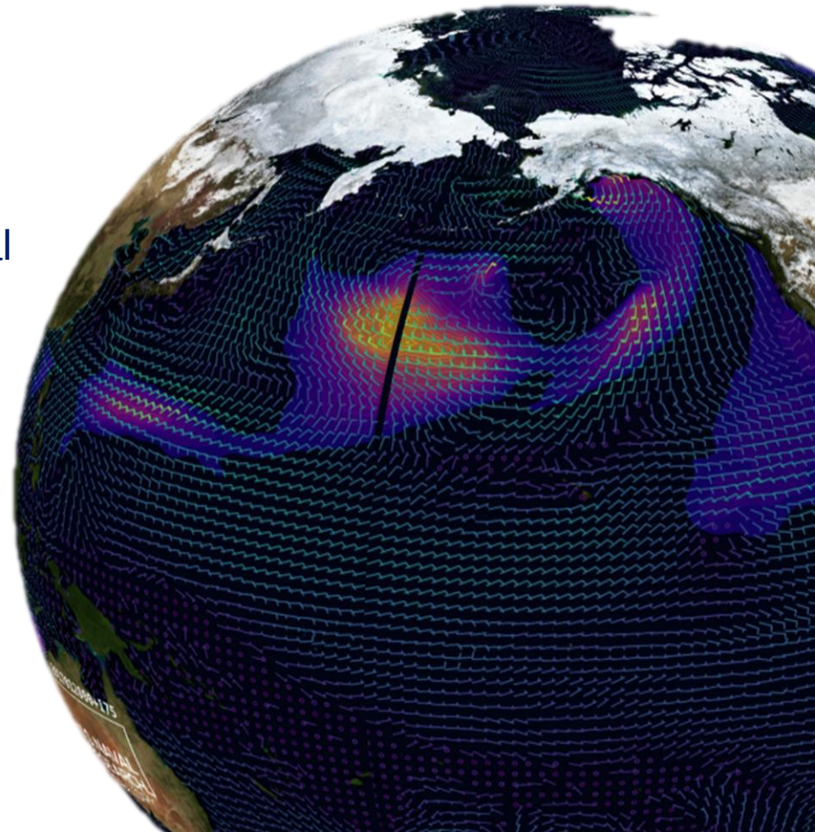
NAVGEM 3.x

- ~13km resolution with new physics-dynamics coupling with NUOPC interoperable physics driver and limited coupling to fit within operational time constraints

2018

NAVGEM 2.x

- ~19km resolution with new scalable infrastructure for improved efficiency and scalability on current DoD computational platforms.
- New TLM/Adjoint models for 4D-Var, FSOI, and sensitivity studies.



Toward Future Navy Modeling

Navy METOC mission, basic and applied research, National ESPC, future coordinated national modeling

2020s

Bridge to NEPTUNE

- Further development, evaluation, and integration of NUOPC/ESMF-compliant community physical parameterization suites and additional coupled components.
- Emphasize performance for next-generation compute platforms
- Committed to partnerships

Focus on coupled modeling and collaborative development. Navy METOC requirements include high-resolution ocean models which form the present and emerging drivers of computational resources for operational Navy earth system prediction.

time constraints

2018

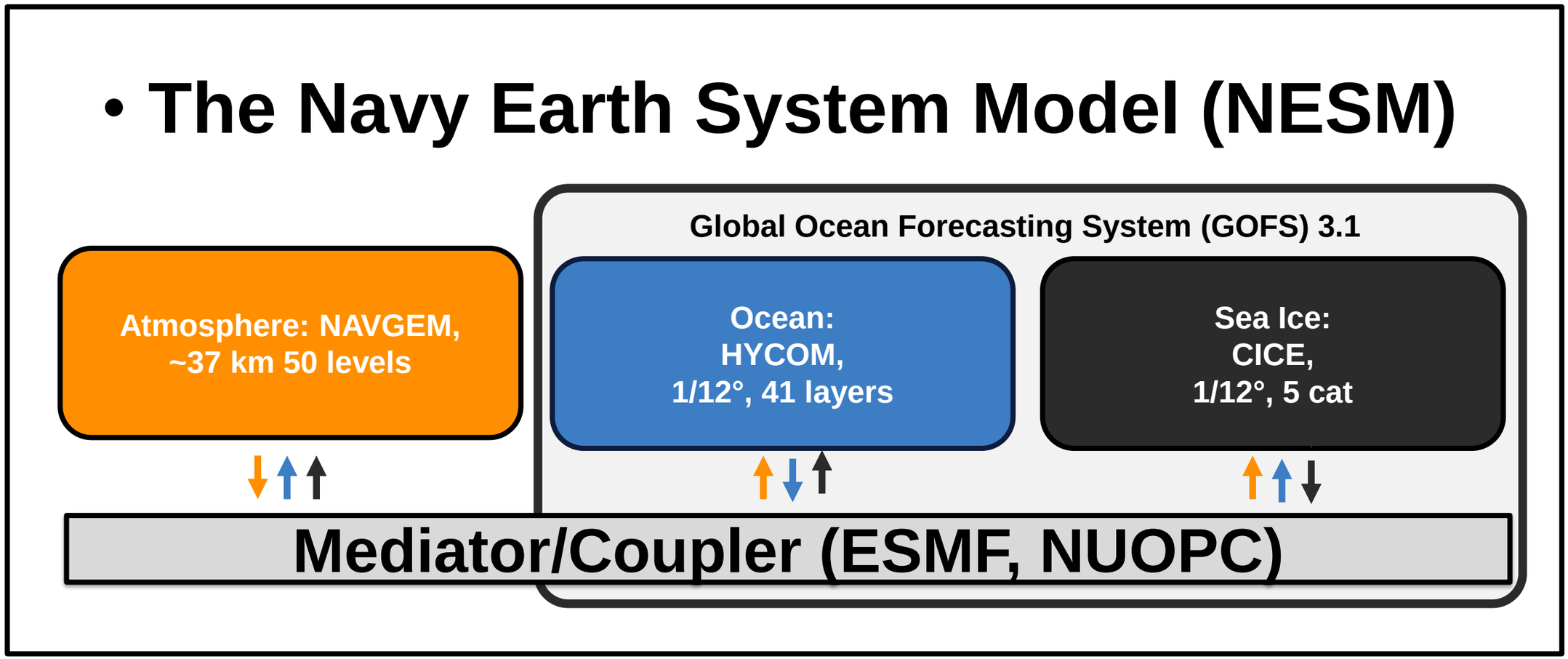
NAVGEM 2.x

- ~19km resolution with new scalable infrastructure for improved efficiency and scalability on current DoD computational platforms.
- New TLM/Adjoint models for 4D-Var, FSOI, and sensitivity studies.

U.S. NAVAL RESEARCH LABORATORY

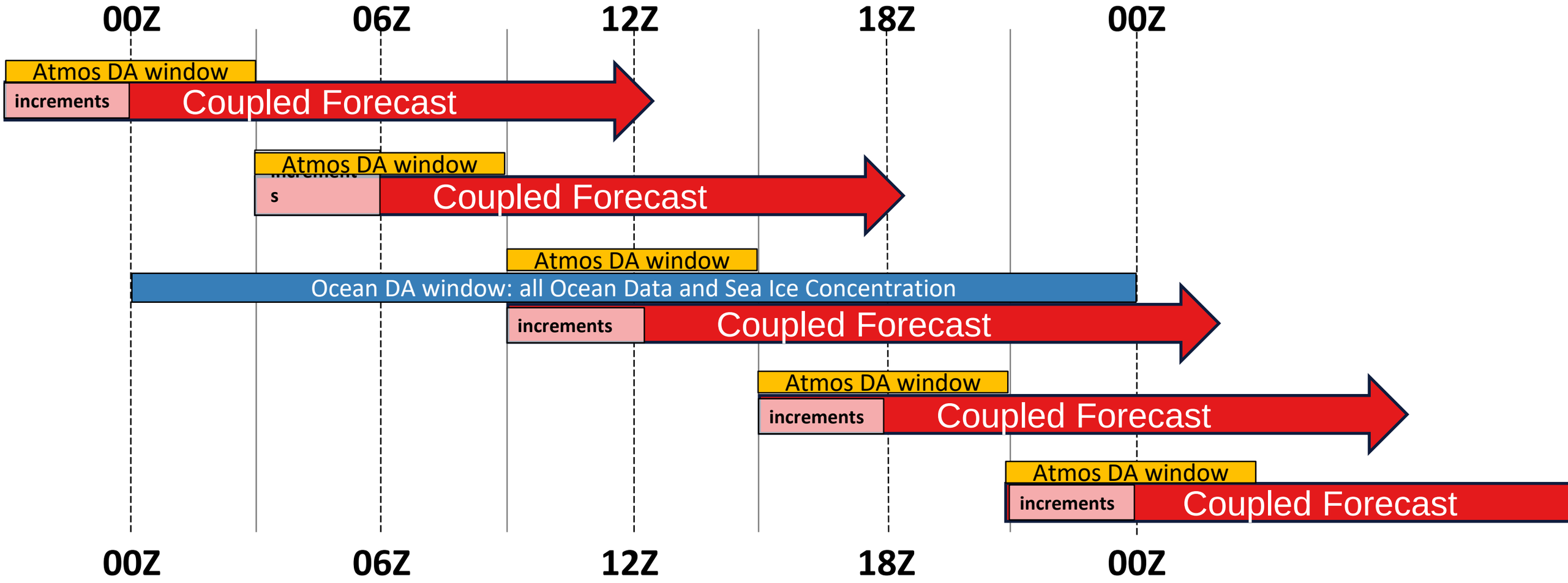
What is NESM?

The Navy's Global Coupled System Based on Current Operational Systems



(N. Barton)

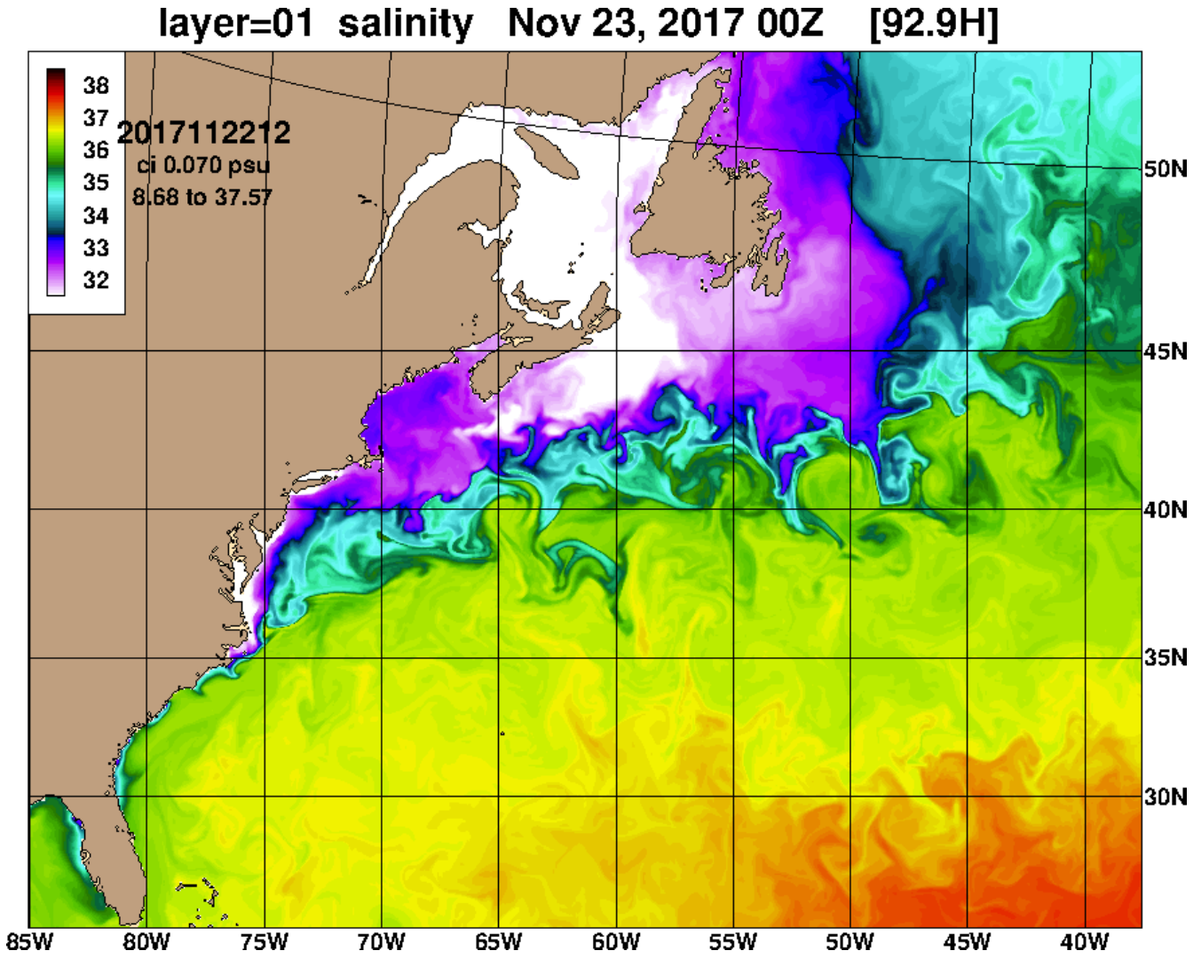
Loosely Coupled DA System Based on Current Systems (NAVDAS-AR/NCODA)

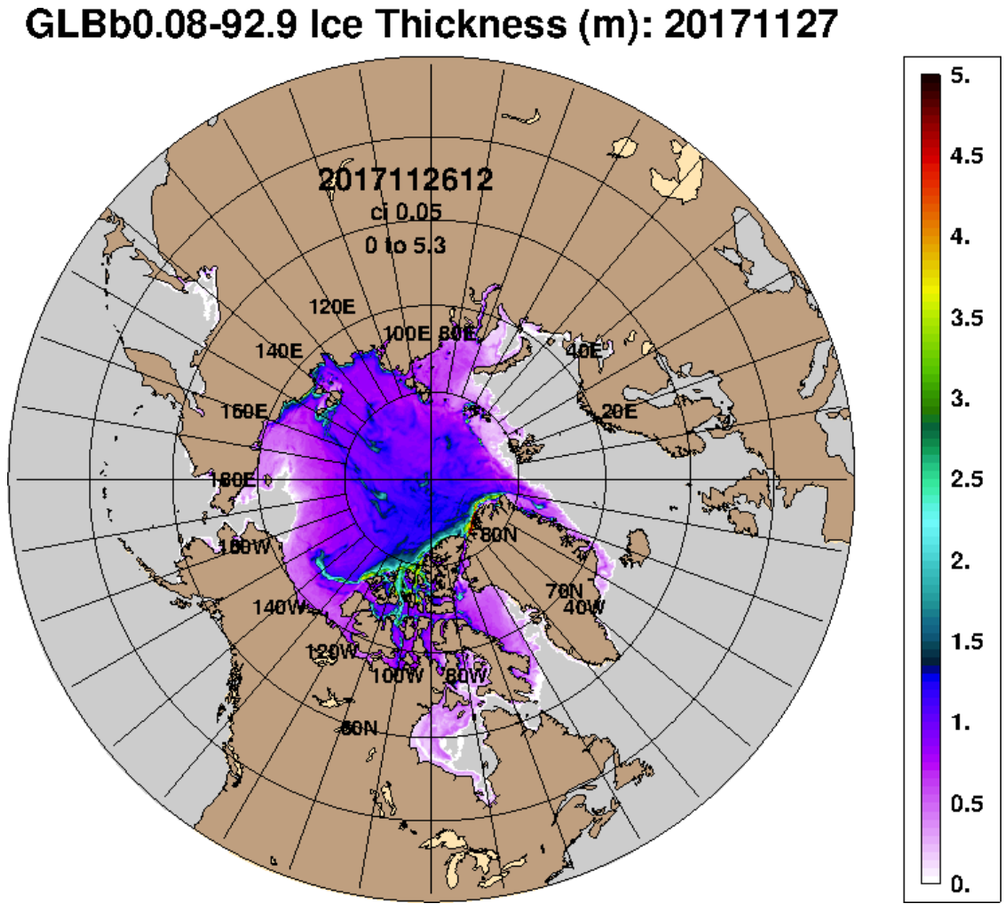


(N. Barton, B. Ruston, J. Metzger)

U.S. NAVAL RESEARCH LABORATORY

Global High Ocean Resolution





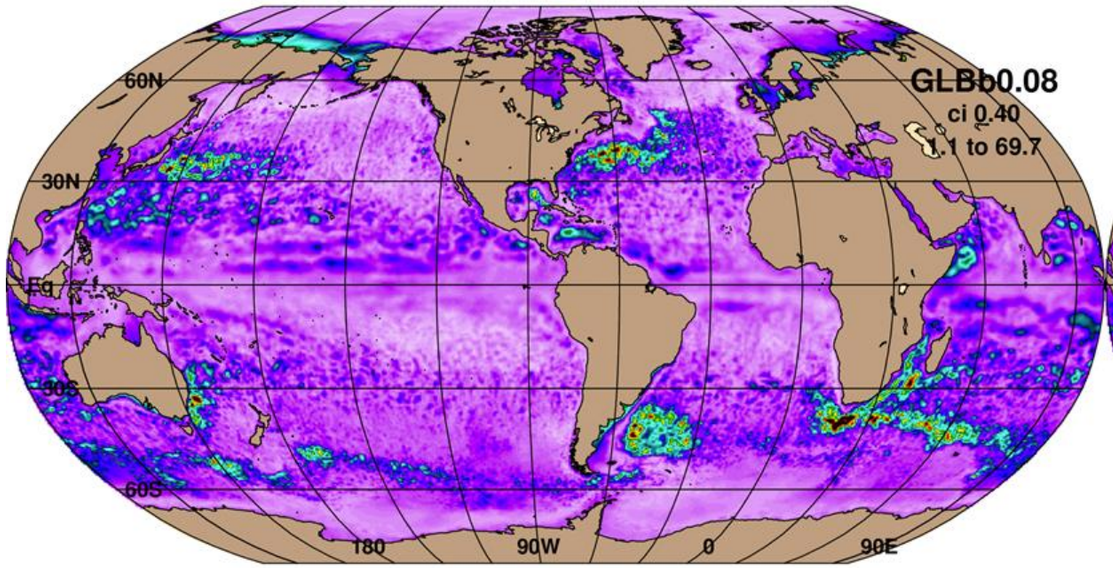
- High fidelity forecasts needed for Atmosphere, Ocean, and Sea Ice

(J. Metzger)

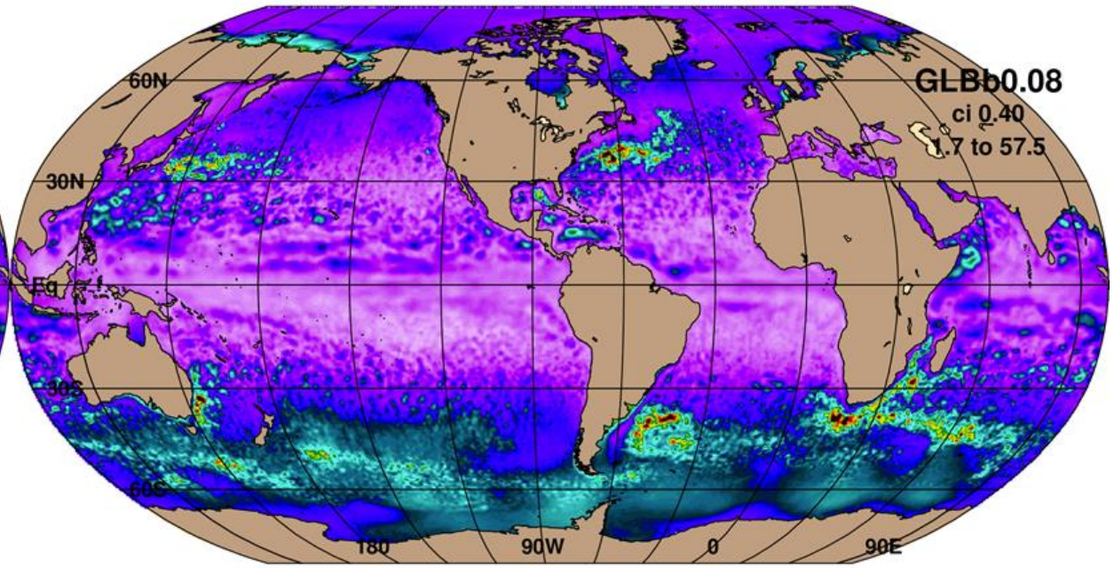
NESM Compared to GOFS 3.1

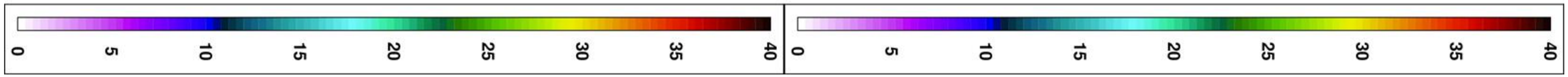
Sea Surface Height (SSH) Variability:

GOFS 3.1



NESM





- **NESM has a dynamic sea surface pressure forcing, while GOFS3.1 does not**

(J. Metzger)

Computational Efficiency

Optimal Core Counts and Load Balancing

- Give operational partners more flexibility when trying to fit the ESPC systems into the operational environment

Ensemble ESPC (single member)				
Total cores	NAVGEM cores	HYCOM cores	CICE cores	Walltime (hours) 45-day forecast
1572	90	1332	150	10.7
1918	107	1631	180	9.4
2227	120	1882	225	8.7

U.S. NAVAL RESEARCH LABORATORY

1 Day of Cycling with Cylc

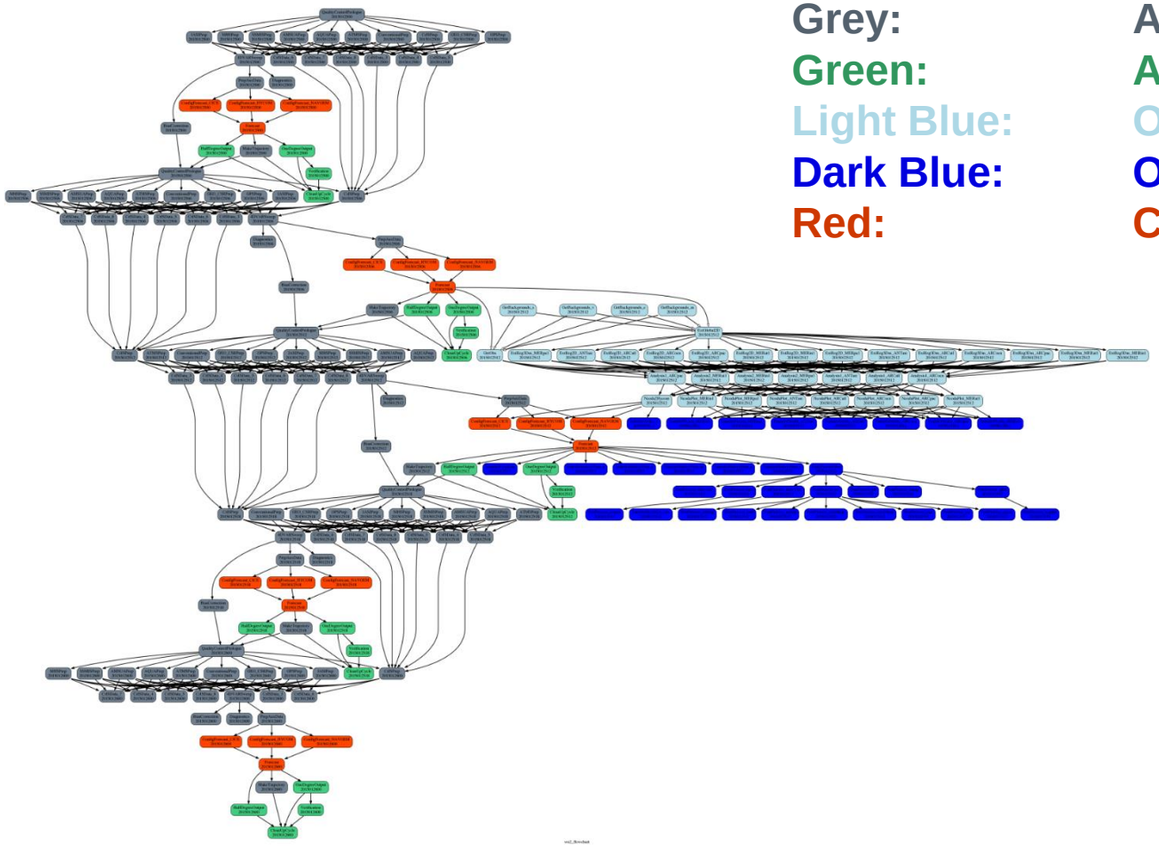
00Z

06Z

12Z

18Z

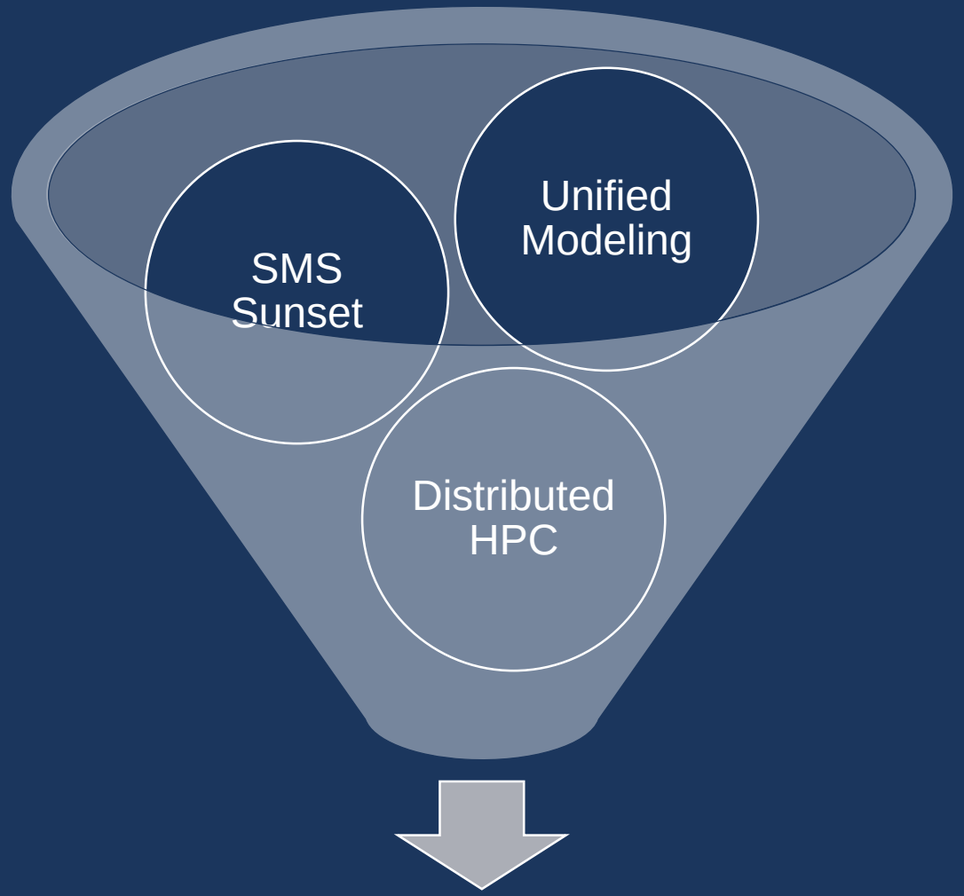
00Z



Grey:	Atmos DA
Green:	Atmos Post Process
Light Blue:	Ocean DA
Dark Blue:	Ocean Post Process
Red:	Coupled Model Tasks

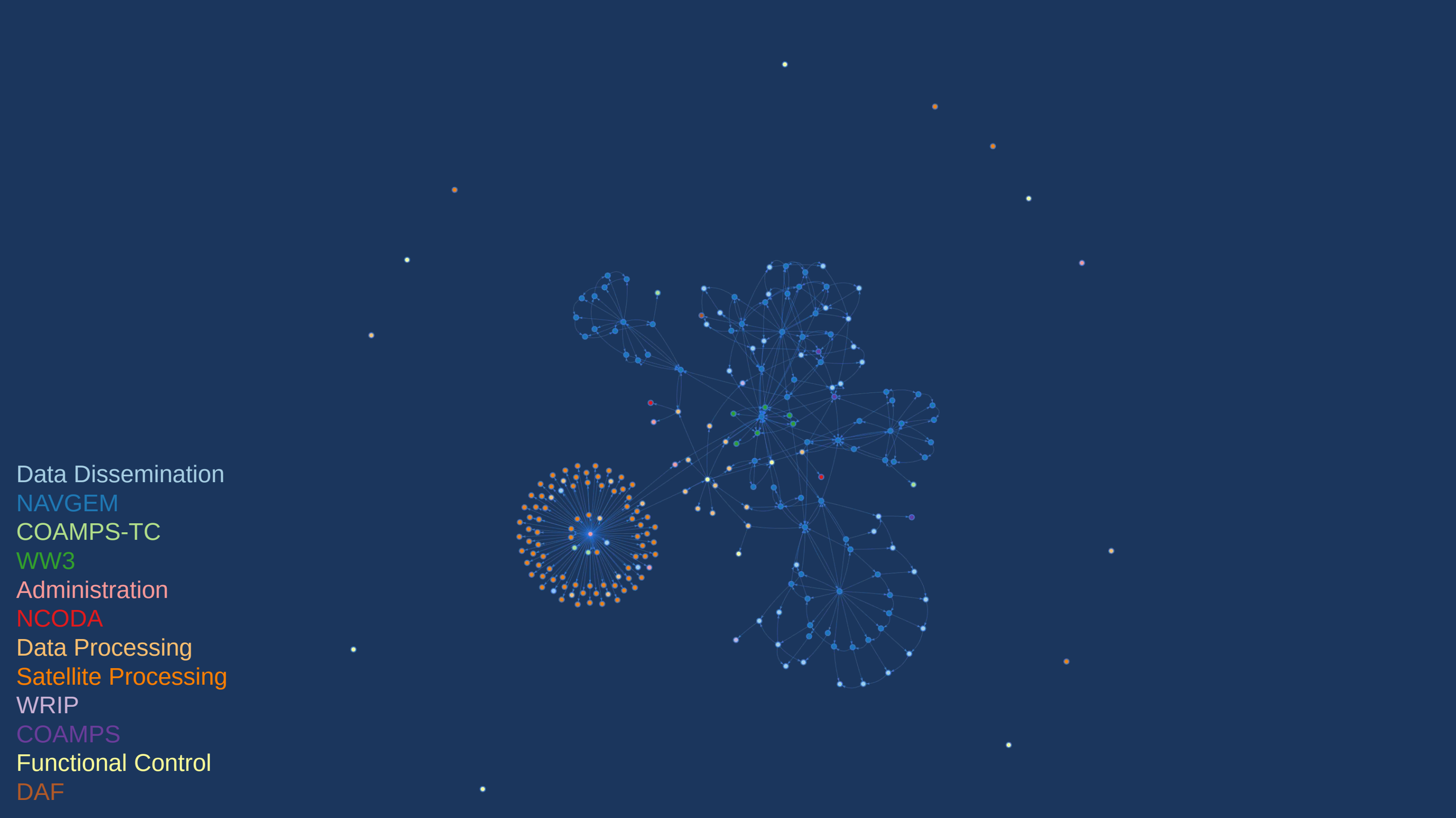
Ensemble is based off perturbed observations, so duplicate this cycle for each ensemble member (!)

Note: Ongoing efforts for modernizing the operational run will result in additional tasks that are more granular.



OPSRUN Modernization

A unique opportunity to examine and re-envision FNMOC operations to reduce errors, provide easier monitoring, leverage new computing capabilities, and control ever-increasing complexity.


Data Dissemination
NAVGEM
COAMPS-TC
WW3
Administration
NCODA
Data Processing
Satellite Processing
WRIP
COAMPS
Functional Control
DAF

U.S. NAVAL RESEARCH LABORATORY

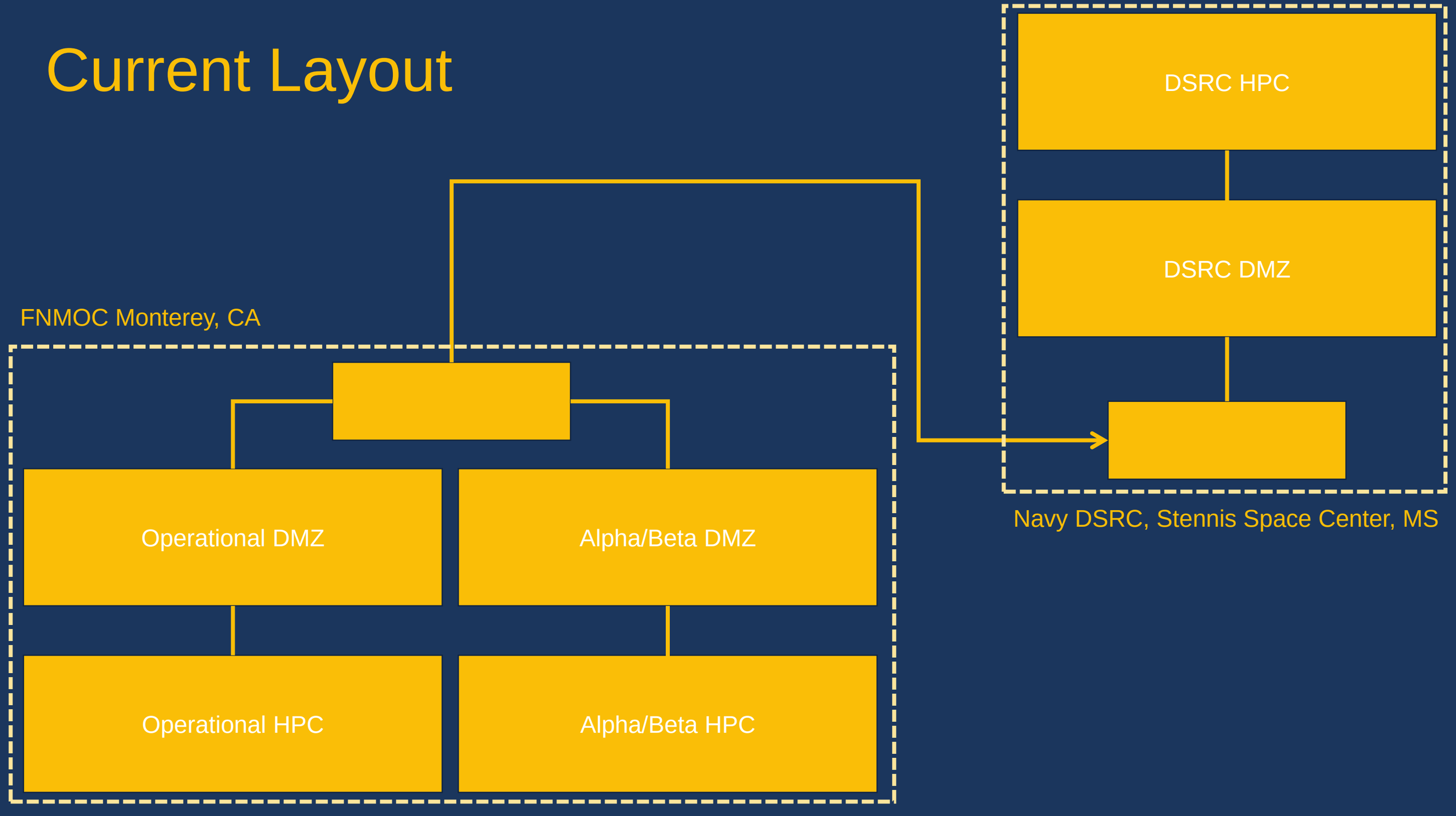
SMS to Cylc Migration

- Porting the existing system to Cylc is suboptimal – seek to *improve*, not just *reproduce*
- Leverage Cylc's unique capabilities and learnings from the existing OPSRUN review
- Cylc developers have now put together a "Suite Design Guide" with best practices

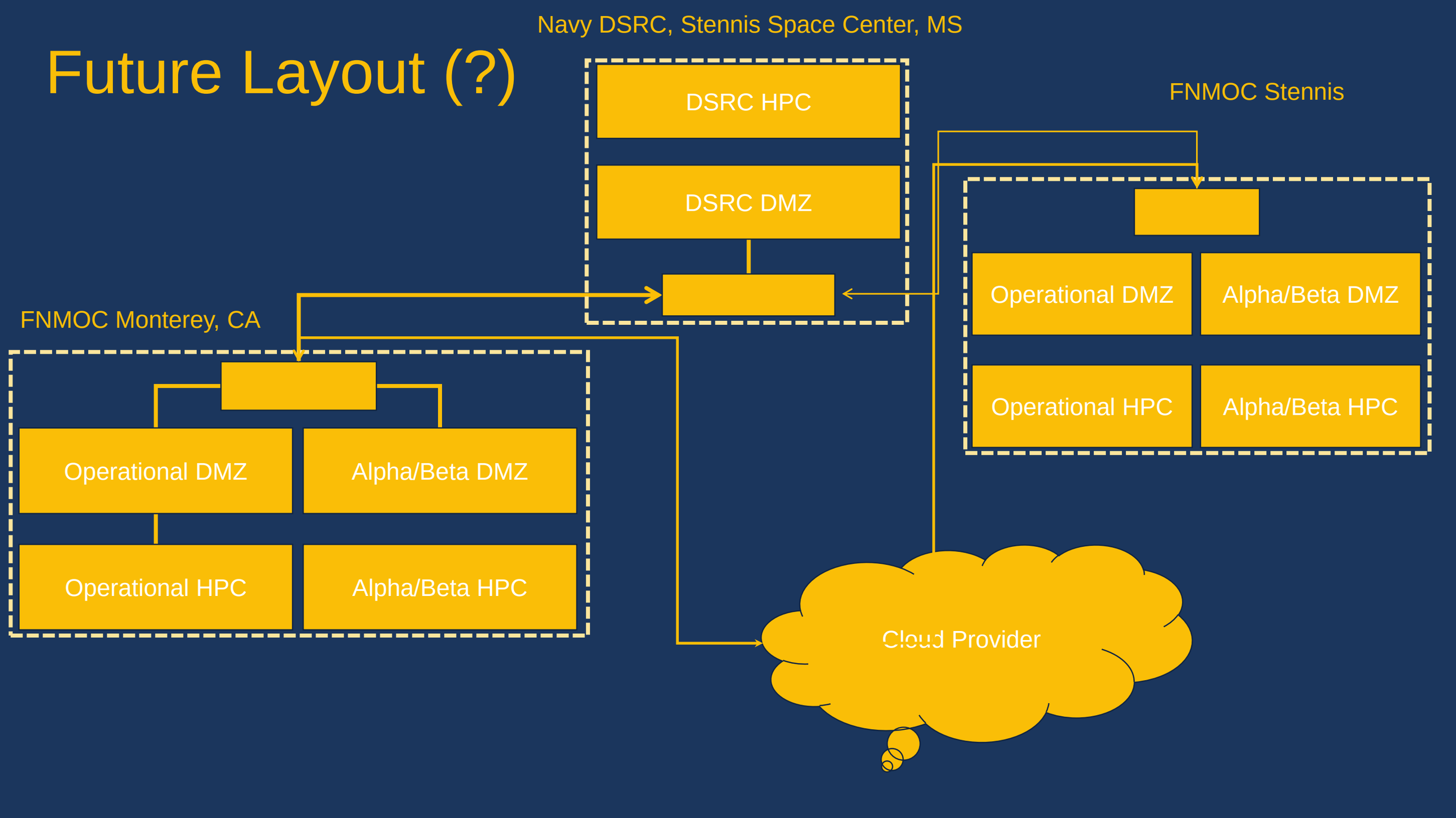
What is done now?
How is it done now?
Why was it done that way?
Is there a different way to accomplish the same thing?
If there is a different way, is it better?

Network Layout

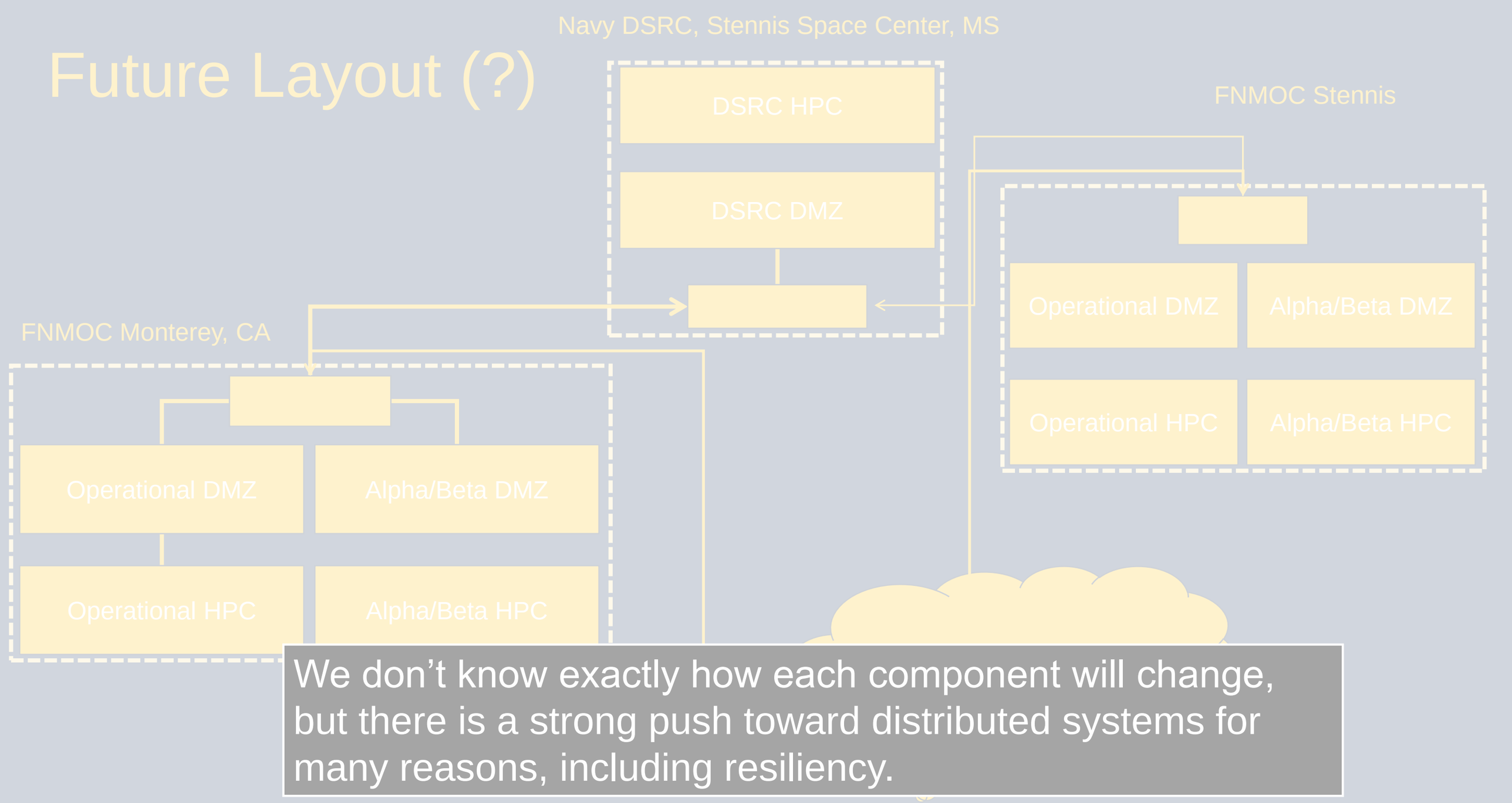
Current Layout



Future Layout (?)



Future Layout (?)



We don’t know exactly how each component will change, but there is a strong push toward distributed systems for many reasons, including resiliency.

U.S. NAVAL RESEARCH LABORATORY

DoD Security Considerations

- Cybersecurity is an immediate and growing concern within U.S. DoD
- Software clearance governed by STIGs (Security Technical Implementation Guide)
  - https://www.stigviewer.com/stig/application_security_and_development/
- Preference for PKI authentication (hard and soft certificates), but username/password currently allowed for machine-to-machine communications.

A Future OPSRUN

Unified Modeling

Unified Monitoring

Fewer errors

Faster error recovery

Increased HPC flexibility

Smoother Transitions

Cloud Computing Evaluations for HPC resource mitigation

U.S. NAVAL RESEARCH LABORATORY

Preliminary

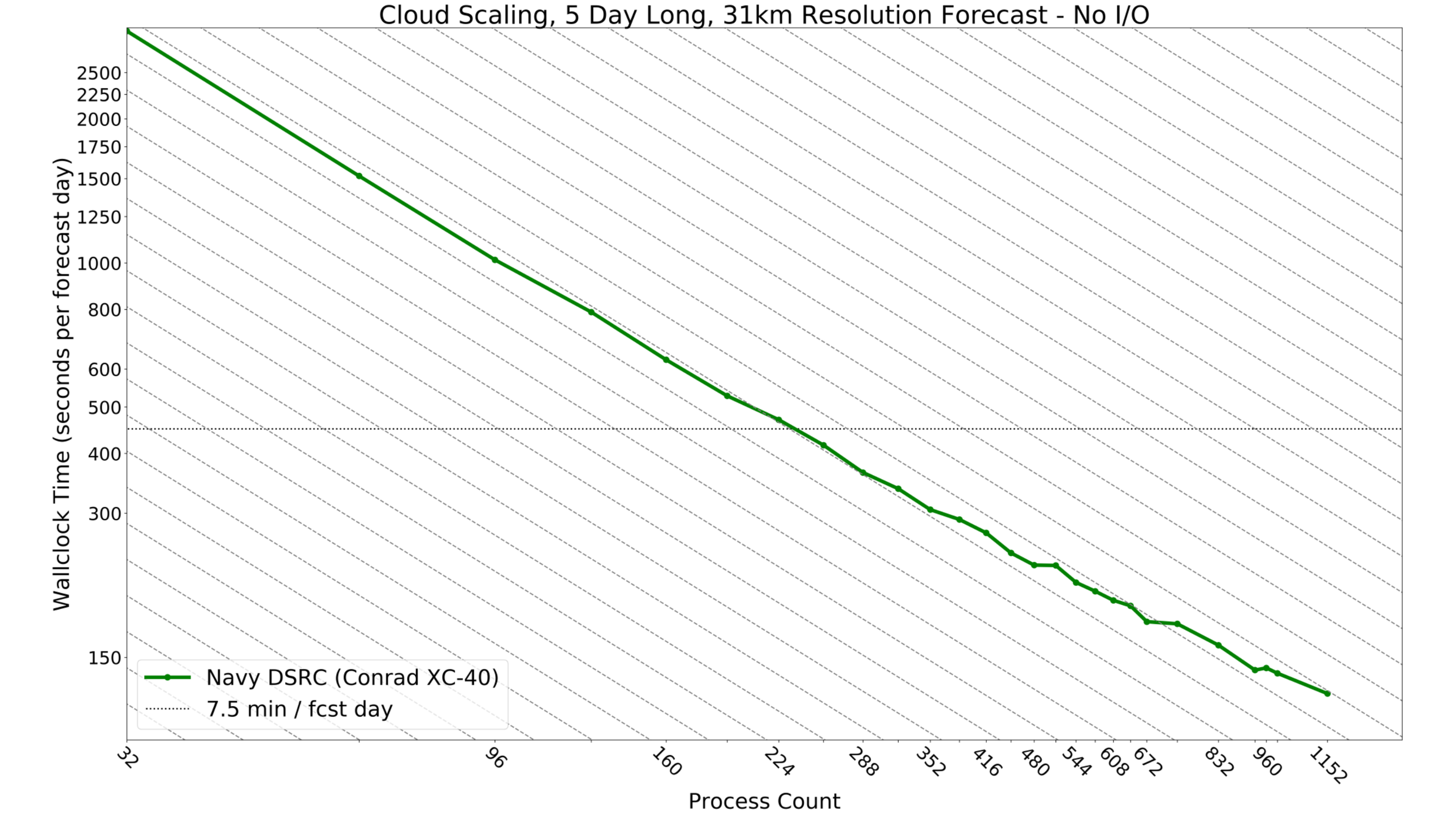
Cloud Computing Evaluations for HPC resource mitigation

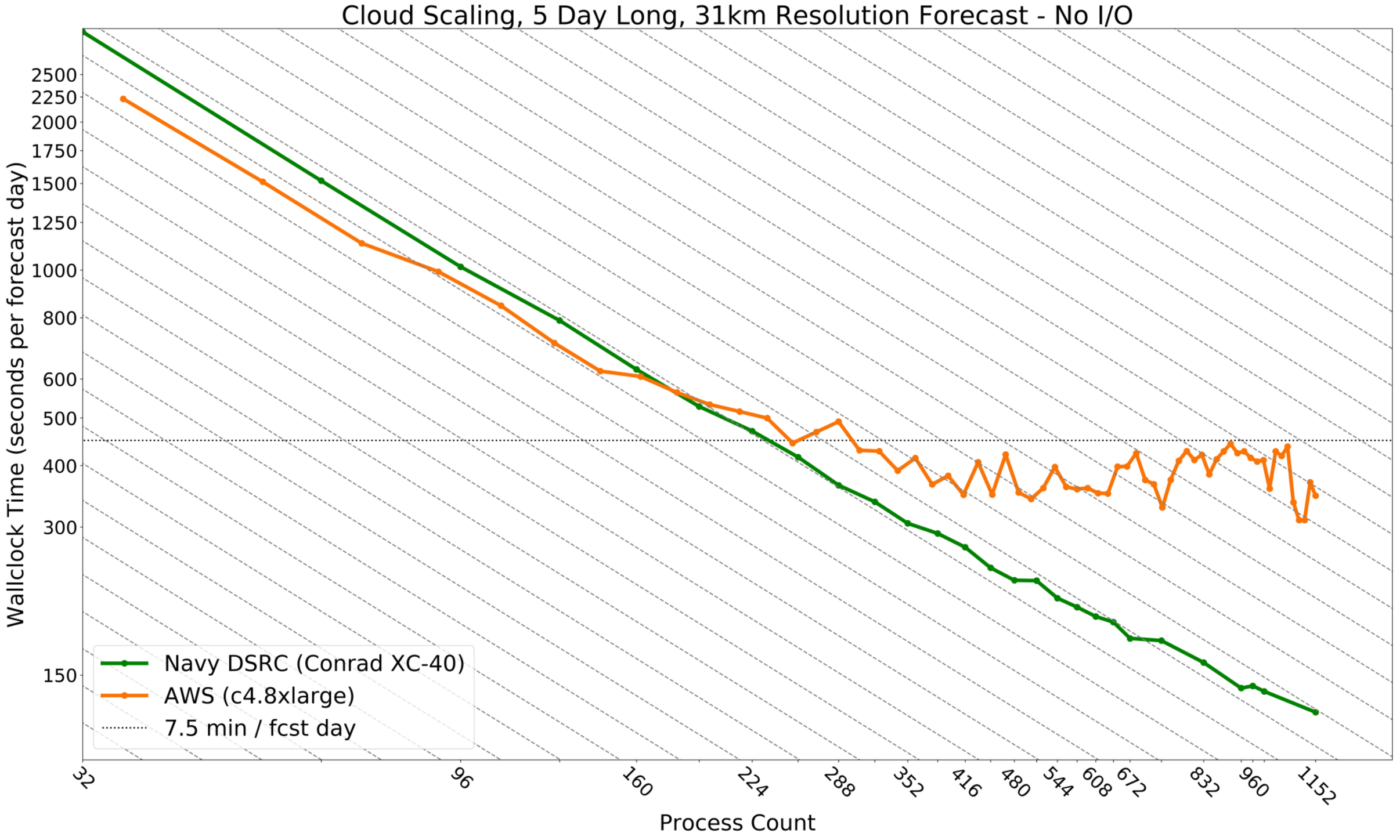
Caveat: focused on computation *only* (i.e. no I/O)

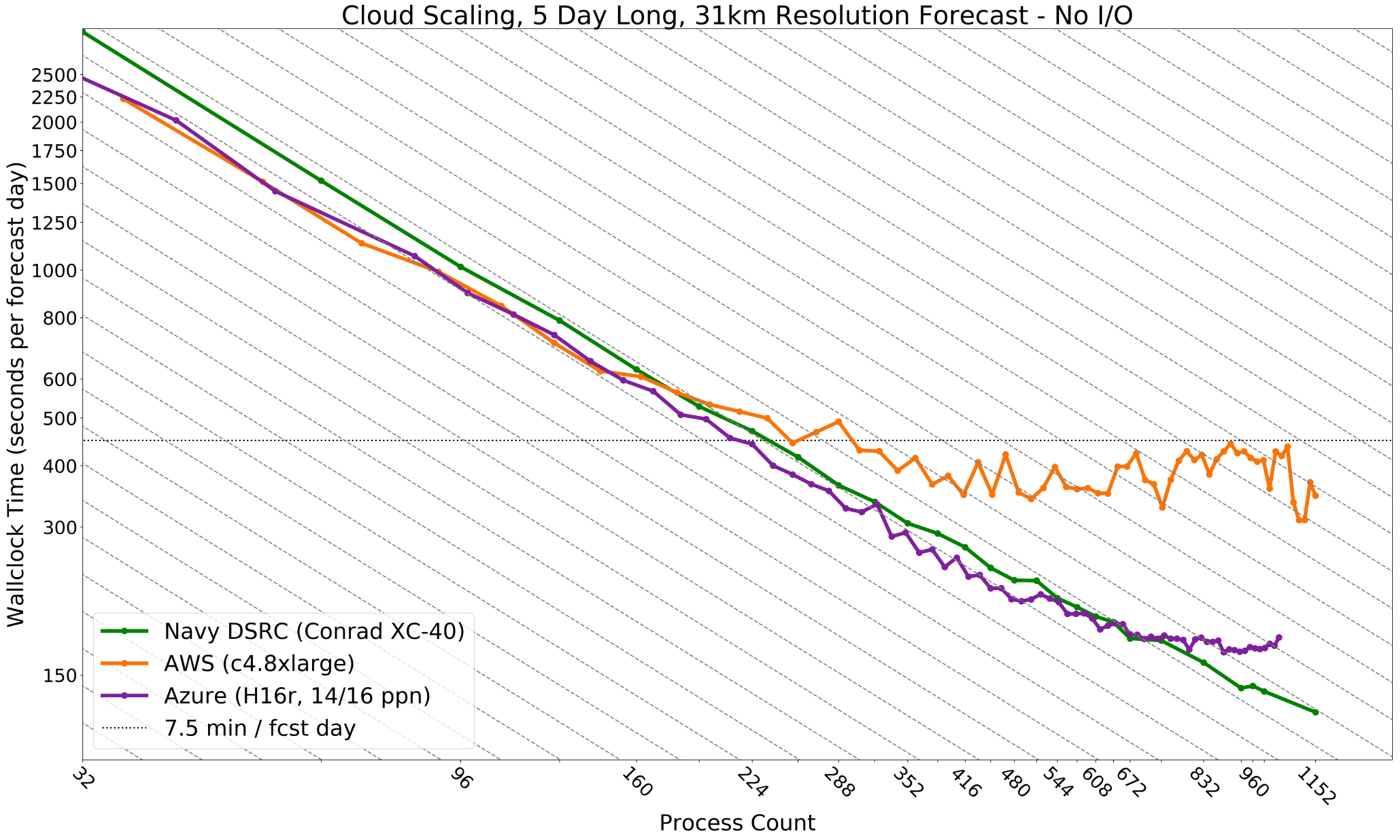
U.S. NAVAL RESEARCH LABORATORY

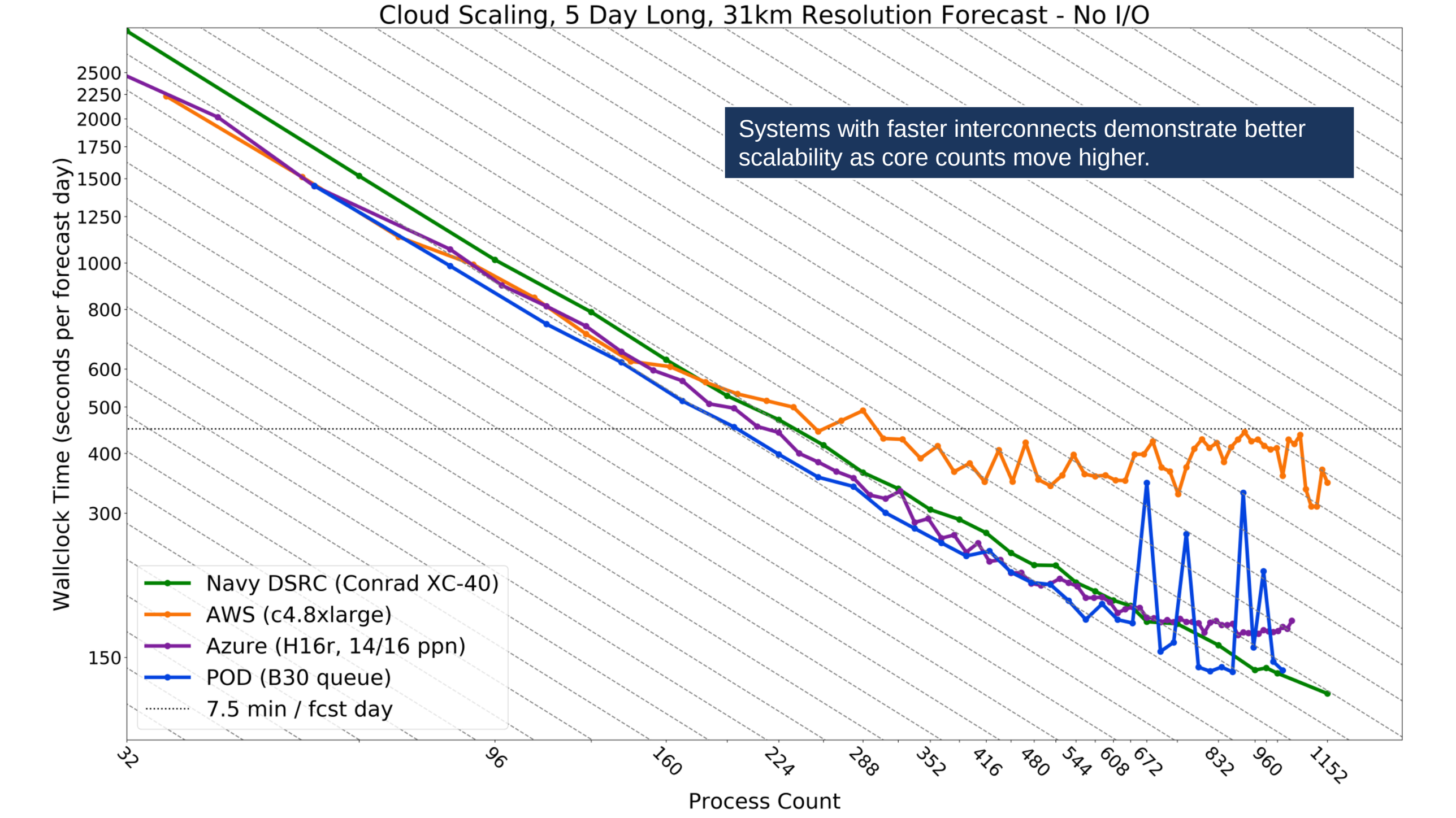
Current Evaluation Platforms

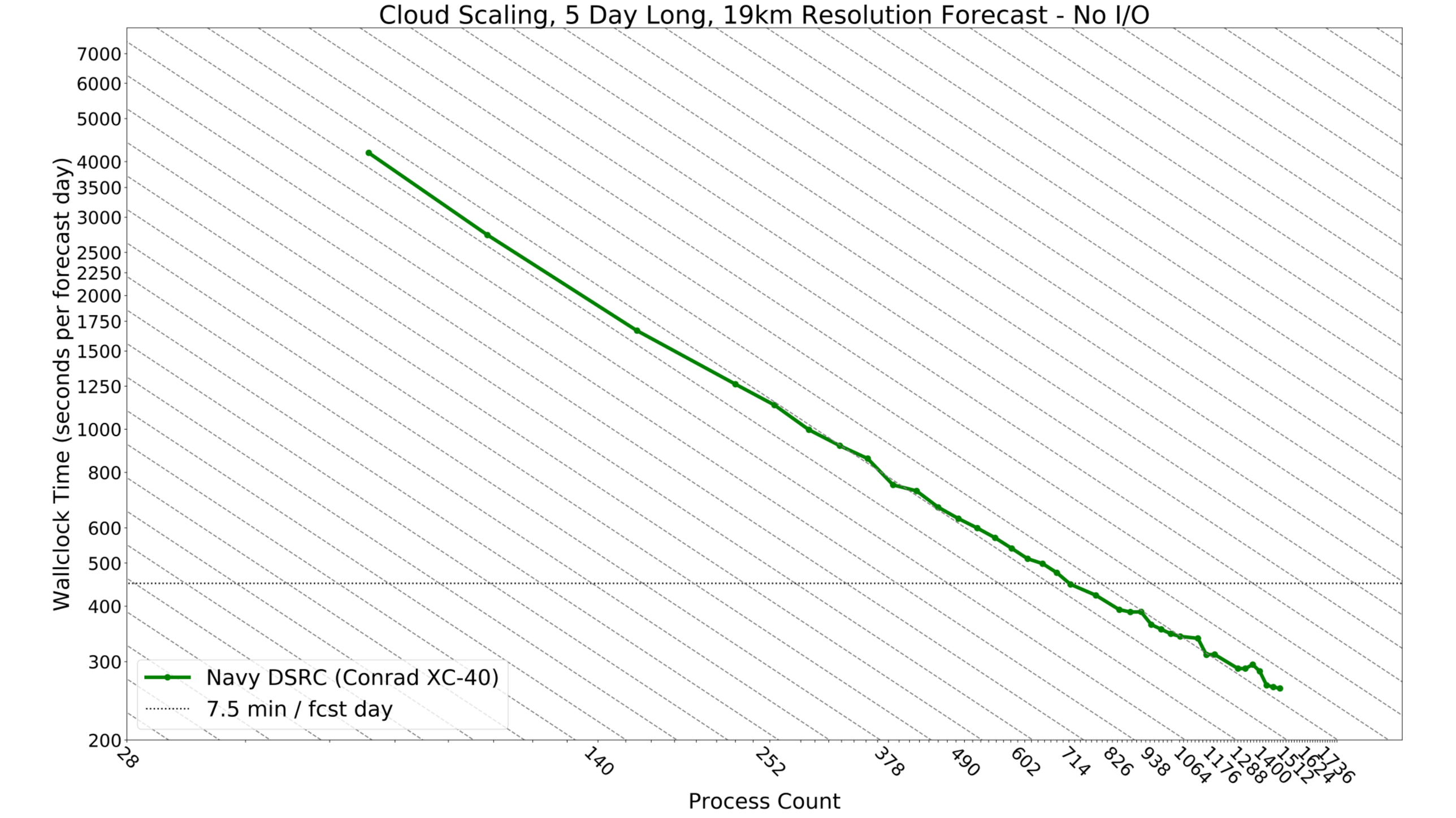
Vendor	Instance Type	Processor	Interconnect
Cray	XC-40	Intel Xeon E5-2698 v3 (Haswell)	Cray Aries/Dragonfly
Amazon	c4.8xlarge	Intel Xeon E5-2666 v3 (Haswell)	Ethernet enhanced with single root I/O virtualization
Microsoft	H16r	Intel Xeon E5-2667 v3 (Haswell)	FDR Infiniband
Penguin On Demand	B30	Intel Xeon E5-2680 v4 (Broadwell)	Intel OmniPath

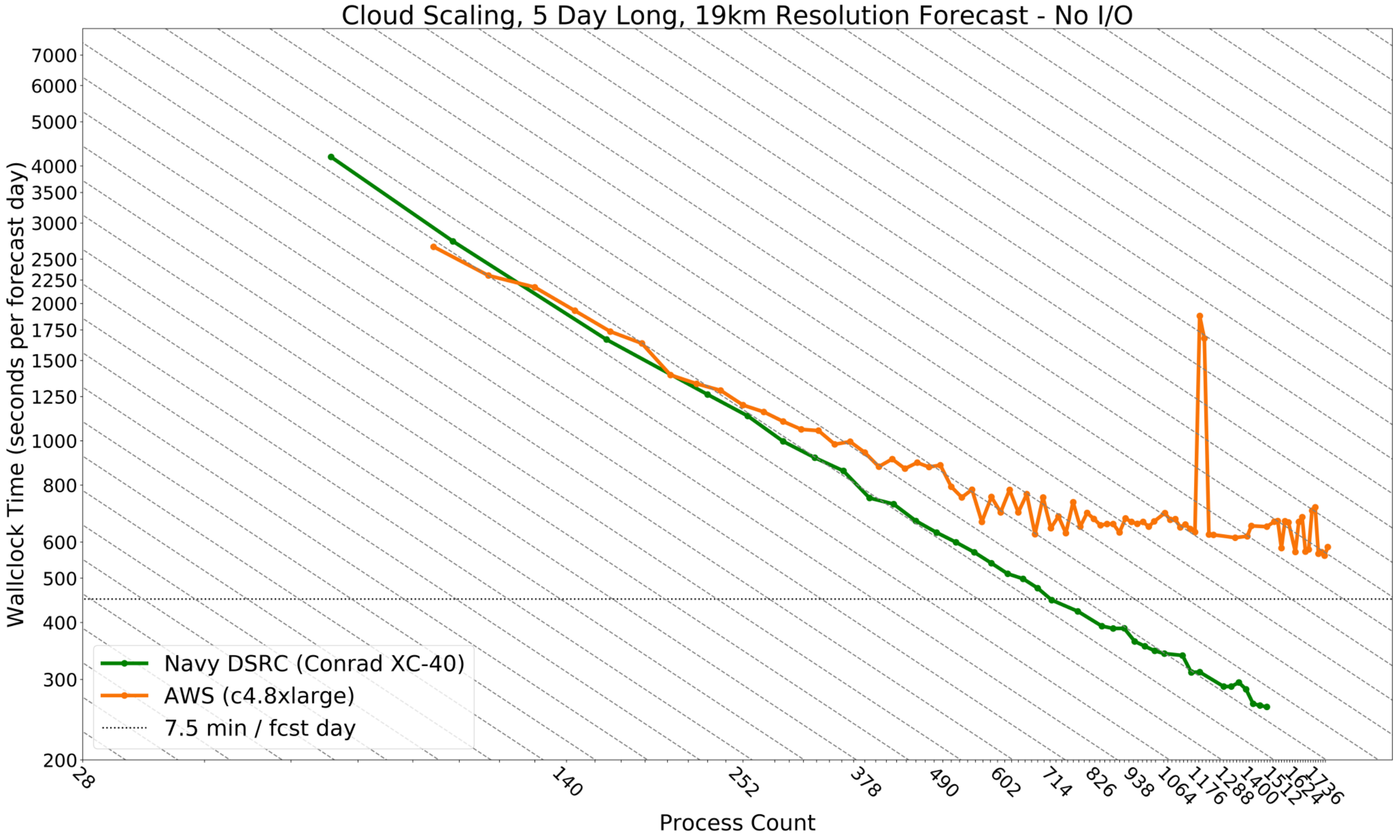

Cloud Scaling, 5 Day Long, 31km Resolution Forecast - No I/O
Wallclock Time (seconds per forecast day)
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
150
Navy DSRC (Conrad XC-40)
7.5 min / fcst day
32
96
160
224
288
352
416
480
544
608
672
832
960
1152
Process Count

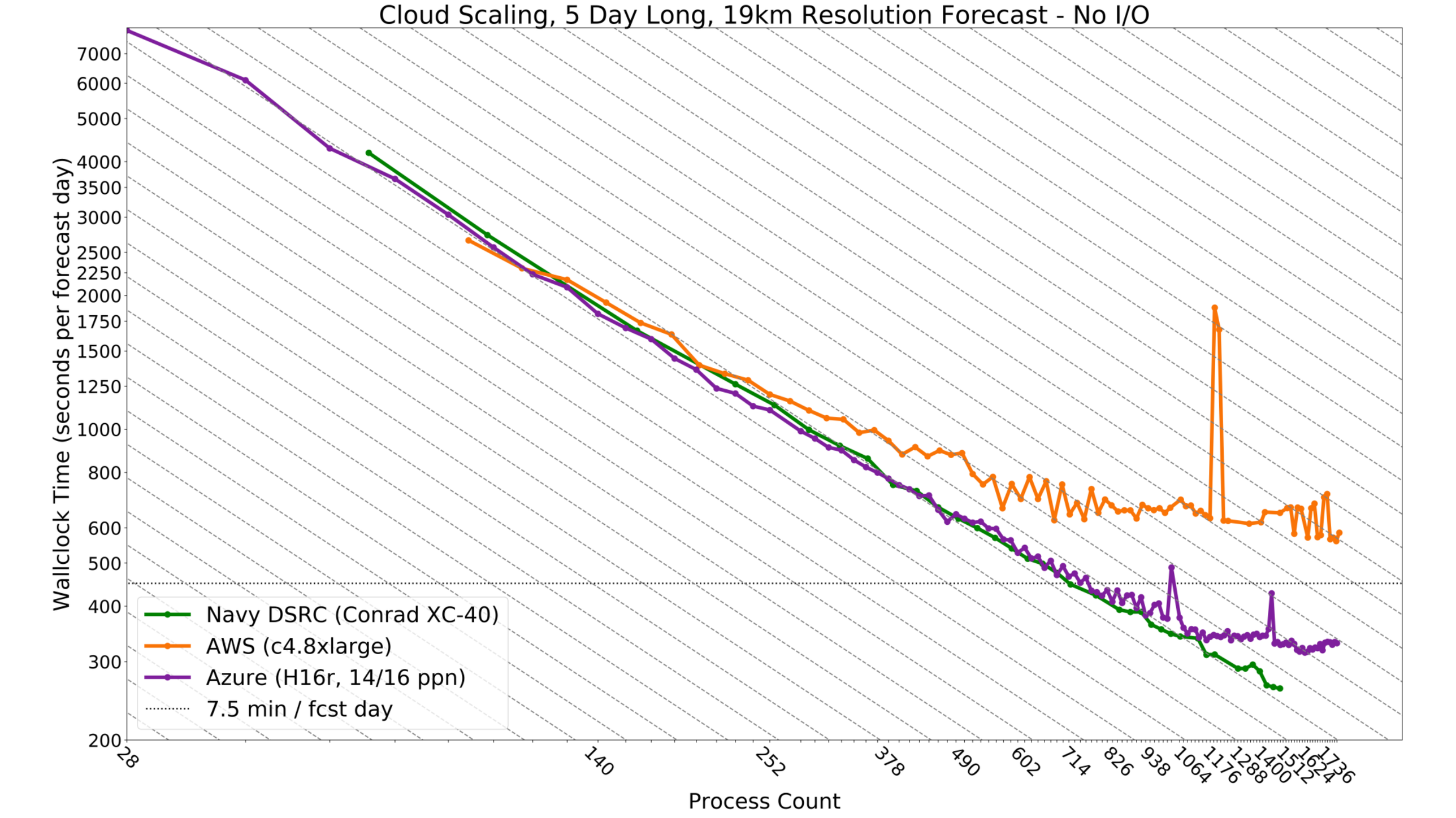

Cloud Scaling, 5 Day Long, 31km Resolution Forecast - No I/O
Wallclock Time (seconds per forecast day)
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
150
Navy DSRC (Conrad XC-40)
AWS (c4.8xlarge)
7.5 min / fcst day
32
96
160
224
288
352
416
480
544
608
672
832
960
1152
Process Count

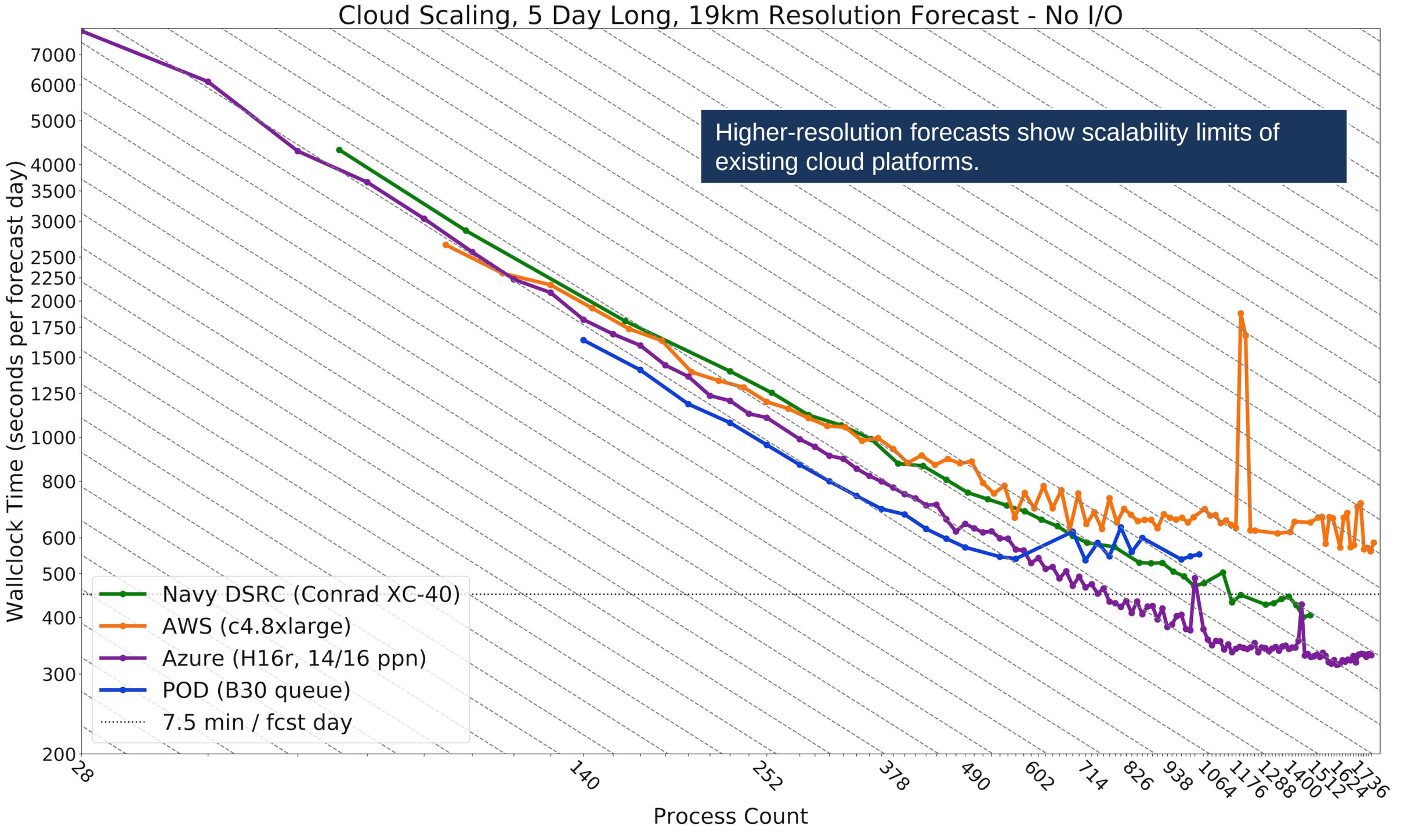

Cloud Scaling, 5 Day Long, 31km Resolution Forecast - No I/O
Wallclock Time (seconds per forecast day)
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
150
Navy DSRC (Conrad XC-40)
AWS (c4.8xlarge)
Azure (H16r, 14/16 ppn)
7.5 min / fcst day
32
96
160
224
288
352
416
480
544
608
672
832
960
1152
Process Count


Cloud Scaling, 5 Day Long, 31km Resolution Forecast - No I/O
Systems with faster interconnects demonstrate better scalability as core counts move higher.
Wallclock Time (seconds per forecast day)
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
150
Navy DSRC (Conrad XC-40)
AWS (c4.8xlarge)
Azure (H16r, 14/16 ppn)
POD (B30 queue)
7.5 min / fcst day
32
96
160
224
288
352
416
480
544
608
672
832
960
1152
Process Count


Cloud Scaling, 5 Day Long, 19km Resolution Forecast - No I/O
Wallclock Time (seconds per forecast day)
7000
6000
5000
4000
3500
3000
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
200
28
140
252
378
490
602
714
826
938
1064
1176
1288
1400
1512
1624
1736
Process Count
Navy DSRC (Conrad XC-40)
7.5 min / fcst day


Cloud Scaling, 5 Day Long, 19km Resolution Forecast - No I/O
Wallclock Time (seconds per forecast day)
7000
6000
5000
4000
3500
3000
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
200
28
140
252
378
490
602
714
826
938
1064
1176
1288
1400
1512
1624
1736
Process Count
Navy DSRC (Conrad XC-40)
AWS (c4.8xlarge)
7.5 min / fcst day


Cloud Scaling, 5 Day Long, 19km Resolution Forecast - No I/O
Wallclock Time (seconds per forecast day)
7000
6000
5000
4000
3500
3000
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
200
28
140
252
378
490
602
714
826
938
1064
1176
1288
1400
1512
1624
1736
Process Count
Navy DSRC (Conrad XC-40)
AWS (c4.8xlarge)
Azure (H16r, 14/16 ppn)
7.5 min / fcst day


Cloud Scaling, 5 Day Long, 19km Resolution Forecast - No I/O
Higher-resolution forecasts show scalability limits of existing cloud platforms.
Wallclock Time (seconds per forecast day)
7000
6000
5000
4000
3500
3000
2500
2250
2000
1750
1500
1250
1000
800
600
500
400
300
200
Navy DSRC (Conrad XC-40)
AWS (c4.8xlarge)
Azure (H16r, 14/16 ppn)
POD (B30 queue)
7.5 min / fcst day
28
140
252
378
490
602
714
826
938
1064
1176
1288
1400
1512
1624
1736
Process Count

U.S. NAVAL RESEARCH LABORATORY

Conclusions

- Coupled high-resolution global probabilistic forecasts will push resource requirements to new boundaries
- Drive for distributed computing coming from multiple levels
- Distributed computing resources for NWP production (and all other ancillary processes) pose new security and management challenges
- Challenges for “Cloud First” strategy vs. requirements for NWP cycling systems
- Different cloud vendors offer different strengths and weaknesses
- The landscape is rapidly changing with vendor offerings and capabilities, so these results will likely be very different at the next workshop