# An Elastic Multiarchitecture Cloud-based High Performance Computing Environment for the Global Forecast System

**Stefan Cecelski PhD**, Christopher Cassidy, Noah Lucas, Brian Summa, Robert Haas, Travis Hartman

January 29, 2024

104<sup>th</sup> Annual AMS Meeting



#### Maxar and cloud-based high-performance computing

- Since 2019, Maxar has been running various numerical weather prediction (NWP) models on cloud-based highperformance computing (HPC) within Amazon Web Services (AWS).
- Maxar's main focus is running real-time production Global Forecast System (GFS) deterministic and ensemble NWP systems.
  - These models are configured as close to model configurations used by the National Centers for Environmental Prediction as possible for specific client use cases.
- Cloud-based HPC within AWS has <u>rapidly evolved</u> since 2019 and Maxar's cloud solutions for NWP have matured with these changes to increase resiliency, boost performance and reduce cost.





#### **Pros of cloud HPC**

- Entirely elastic: can spin up/down all resources to only use costly resources when required (servers, file systems, etc.)
- Up-to-date technology and server flexibility: leverage different and/or newer servers and technologies as they are released; not stuck on aging hardware
- Resiliency, redundancy and geographic diversity: multiple cloud-based and clusterbased strategies to "fail fast"

# Cons of cloud HPC

 Cost: solutions need to be well architected to mitigate overspending

- Black box: not *every* cloud resource is highly controllable/configurable and for finetuning HPC for NWP, having control is desirable
- Debugging and mitigating problems is more complex: a lot more moving parts and lack of visibility due to being "black box"



# A fully automated and cost-effective workflow

- The entire workflow, from ingesting data, to creating the cloud HPC cluster, and finally relaying the forecast data, is fully automated using various AWS services.
- Costly components, such as the cloud HPC cluster, are reallocated each forecast cycle to only incur costs while in use. On Cluster Cluster **Process** Stack This workflow is deployed to several AWS regions using continuous deployment strategies.





# Maxar's journey of cloud-based HPC for NWP



INTELLIGENC



INTELLIGENCE

### Learning experience #1: Spin up and down resources

![](_page_6_Figure_2.jpeg)

- Do not take a "spin up and forget it" approach to configuring and building your cloud HPC cluster: they are expensive.
- Instead, design repeatable cloud-based workflows that can spin up and down your cloud HPC resources when you need them.
- Infrastructure planning ahead of time will save a lot of future costs given the number of resources and complexity of HPC clusters.

![](_page_6_Picture_6.jpeg)

### Learning experience #2: Cluster architecture flexibility

- Cloud HPC enables the flexibility of an evergrowing number of server/instance/etc. types and sizes, take advantage of them and create "recipes" for each!
- This flexibility can help mitigate...
  - Capacity concerns (not enough servers in a specific region and/or zone)
  - Performance and cost issues (don't get stuck on aging architecture that is slower and costs more!)

![](_page_7_Figure_5.jpeg)

![](_page_7_Picture_6.jpeg)

#### Learning experience #3: Fail fast

- For production cloud HPC workloads like real-time NWP forecast generation, being able to pivot to new resources/zones/regions upon failure in a time efficient manner is critical.
- Examples of failing fast on AWS architecture:
  - "Dueling clusters": more than one cluster running the same workload in multiple regions or zones (quite costly)
  - Capacity reservations: if needing a very large cluster, reserve capacity of your preferred server type in your preferred region; if not available, pivot to other architectures or regions (see lesson #2).
  - On-cluster healing: implement tactics to handle bad servers/jobs gracefully; AWS HPC tools can automatically swap out bad servers as detected.
  - Advanced monitoring: since cloud resources are more dynamic, keeping a heartbeat of workflows and resources is essential.

![](_page_8_Picture_8.jpeg)

# Wrapping up

- Maxar has built and iterated upon its award-winning cloud HPC environment by using architecture flexibility and "failing fast" as keystones for continuous improvement.
- Using the lessons learned, Maxar has not only improved the performance of its NWP workloads, but also has reduced computational footprint and cost in the process.

Maxar wins AWS Partner Award for Best HPC Solution for Public Sector

![](_page_9_Figure_5.jpeg)

HPC

O Harris

O Sectors

O Technologie

0 COVID-18

AIMUDE

C Special

Basevro

C Codeas

C Frank

0 .00 Dan

O Our Author

Maxar Blog: http://bit.ly/Maxar-HPC

![](_page_9_Picture_7.jpeg)

#### AWS 2023 Blog:

https://aws.amazon.com/blogs/hpc/how-maxar-builds-short-duration-bursty-hpcworkloads-on-aws-at-scale/

![](_page_9_Picture_10.jpeg)

![](_page_10_Picture_0.jpeg)

MAXAR.COM

© 2024 Maxar Intelligence Public