PUBLIC SECTOR SUMMIT ONLINE

Forecasting the weather with AWS Cloud

Stefan Cecelski, PhD

Senior Data Scientist and Engineer

Maxar

aws PUBLIC SECTOR SUMMIT ONLINE

Defining Numerical Weather Prediction (NWP)

- NWP is the process of creating representations of future atmospheric states (forecasts) using mathematical models.
  - Fundamental equations of motion, thermo, physics, calculus, chemistry, etc. for every x, y, z, and t around the world + even a random number generator every now and then!

- Several governments and institutions run global (and regional) NWP models to produce forecasts.
  - NOAA → GFS or American
  - ECMWF → Euro or European
  - And others...

Maxar Goals for Running NWP on Cloud HPC

- **Goal #1:** To run and produce forecasts of the Finite Volume Cubed Sphere Global Forecast System (FV3GFS) **faster** than NOAA does for delivery to a number of Maxar clients.
  - Provide clients with the ability to make informed decisions faster based on weather forecasts of 1-15 day lead times.
  - Create custom output (maps and data) that is aligned with customers' business, mission, and operations.

Maxar Goals for Running NWP on Cloud HPC

- **Goal #2:** Produce **highly correlated** forecast output (such as weather patterns and deviations from climatology) to what NOAA produces.

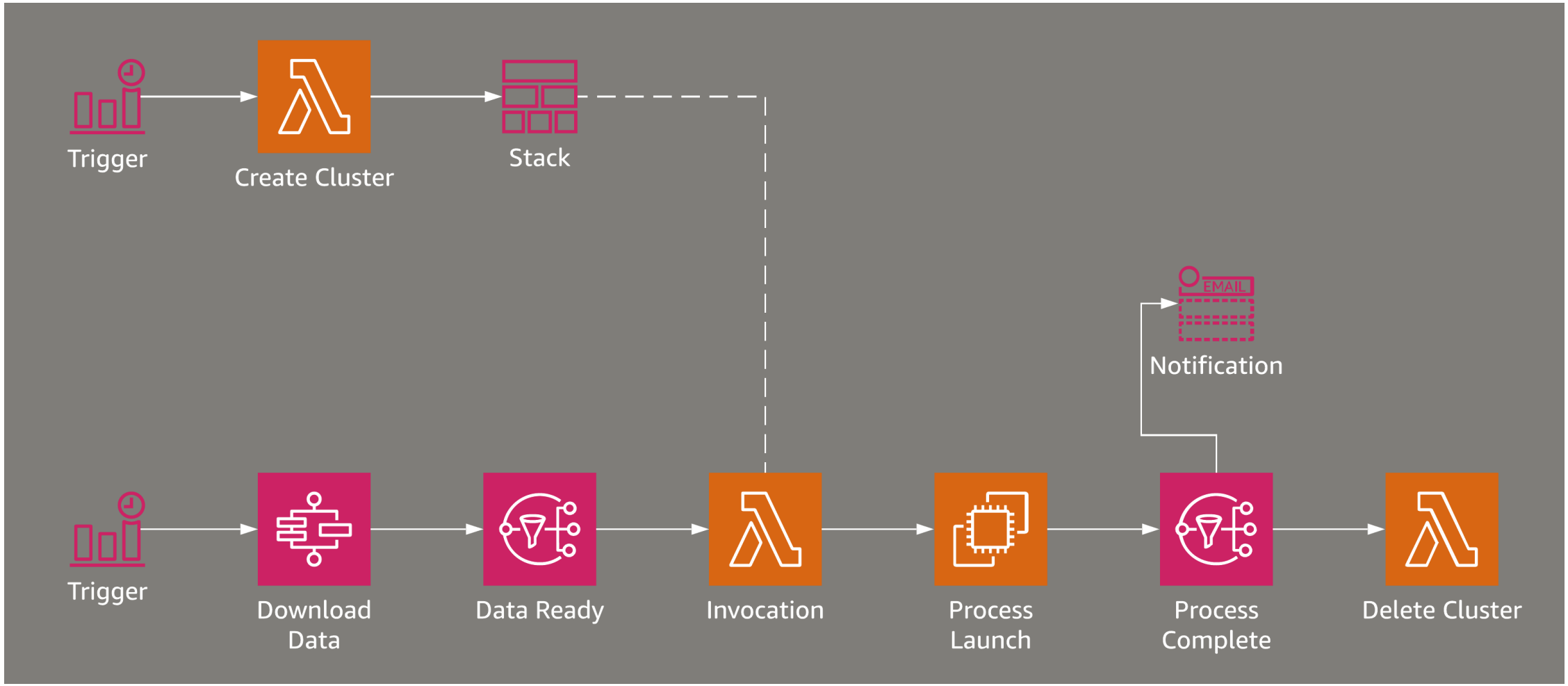
- **Goal #3:** Leverage cloud high performance computing (HPC) built within AWS to **emulate and exceed** the performance of NOAA bare-metal HPC cluster to **produce and deliver** the FV3GFS forecast output.

How Maxar Leverages AWS Cloud HPC for NWP

A Fully Automated Workflow

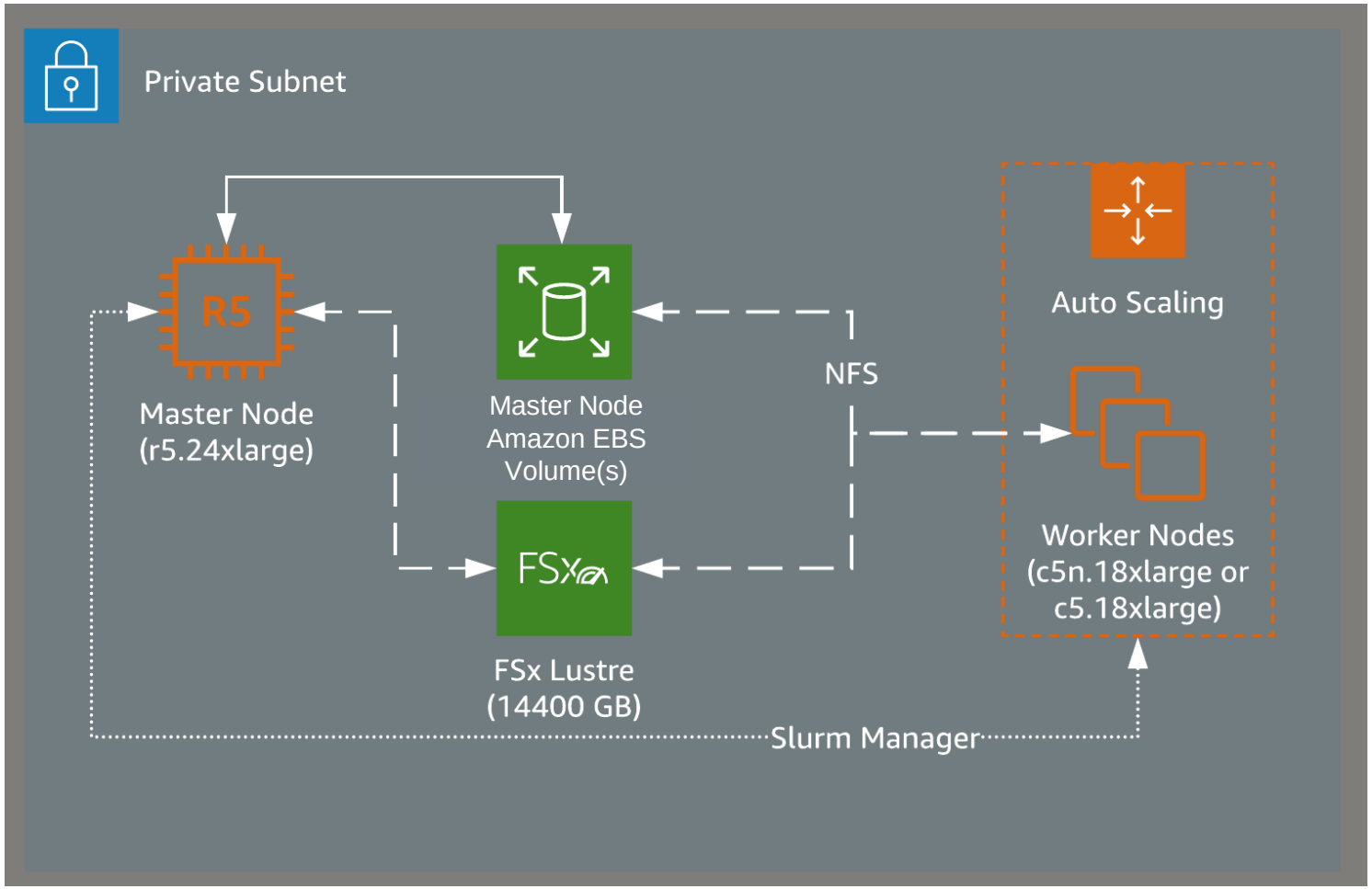
- HPC cluster resources (Amazon EC2 instances, Amazon EBS volumes, FSx Lustre) are re-allocated each day.
- Workflow execution is automated using a series of Step Functions and Lambda functions with AWS SNS notifications.
  - Use a Lambda layer for AWS ParallelCluster to simplify deployment.
  - AWS CloudFormation stacks to build the AWS resources for the workflow.
  - Infrastructure updates are automatically deployed to AWS Regions using CodePipeline.

A Fully Automated Workflow



A Cloud HPC Environment for the FV3GFS Application

- AWS ParallelCluster software orchestrates the spin-up/spin-down of cluster resources.
- FV3GFS is compiled using Intel's Parallel Studio XE compiler environment, 2020 update 1.
- Executed on AmazonLinux 2 operating system.
- Network interconnect is either Elastic Fabric Adapter (EFA) or TCP depending on instance type.



Benchmarking Cluster Performance for NWP

Finding the Ideal Cluster Configuration

- **37** different cloud HPC cluster configurations were built to test the FV3GFS application on as little as **252** cores to over **11,000** cores.

- Maxar currently runs the application in production on both c5.18xlarge and c5n.18xlarge Amazon EC2 instances.
  - **24** of the test cluster configurations utilize c5n.18xlarge with the EFA networking adapter (max 100Gbps).
  - The remaining **13** test configurations use c5.18xlarge instances with TCP networking (max 25Gbps).

- All configurations leverage a **14TB** FSx Lustre file system with a progressive file layout (PFL) across the 12 object storage targets (OSTs).

Assessing FV3GFS Performance

- While numerous metrics can be analyzed to assess the performance of the FV3GFS application on AWS Cloud HPC, we decided to focus on **total wall clock time** initially given its importance for the previously set goals.
  - Other metrics such as integration time, I/O timing, scaling curves, and so forth are also analyzed but not as critically.

- Maxar defines **total wall clock time** as the time it takes the FV3GFS application to complete a 16-day simulation.

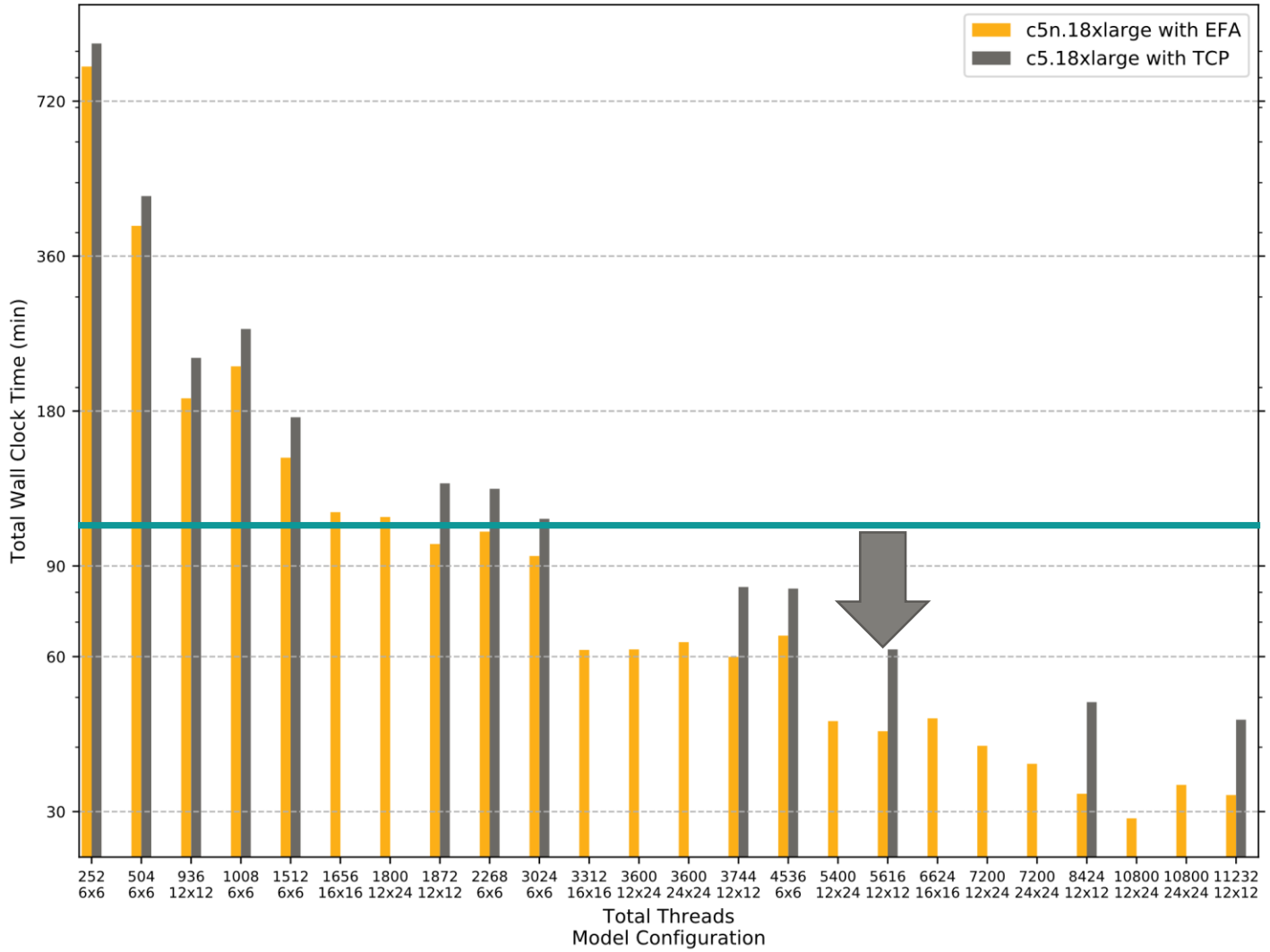
Cirrus performance baselines – 16-day weather forecast

Cluster	Wall Clock Time for 16-day FV3GFS Forecast	Total Node Count	Total Core Count	Number of MPI Threads	Number of OpenMP Threads	Worker Node Amazon EC2 Instance Type	Physical Cores per Instance	RAM per Instance (GB)	Network Max Bandwidth (Gbps)
Cirrus Prod	**~40-43 min**	**156**	**5616**	**936**	**6**	**c5n.18xlarge**	**36**	**192**	**100**
Cirrus Backup	~48-53 min	234	8424	936	9	c5.18xlarge	36	144	25
NOAA*	~100 min	148	3552	1776	2	-	24	-	100

*Estimated configuration based on documentation

Results – FV3GFS total wall clock time

- Numerous configurations outperform NOAA's bare metal time of ~100min (horizontal green line).
- Current production (arrow) configuration is exceptionally performant at its given core count.
  - Other configurations with higher core counts perform worse.
- One configuration achieved a sub-30 minute total wall clock time.
- In general, the 11,000 core count mark coincided with near 30-minute total wall clock times for a 16-day forecast.



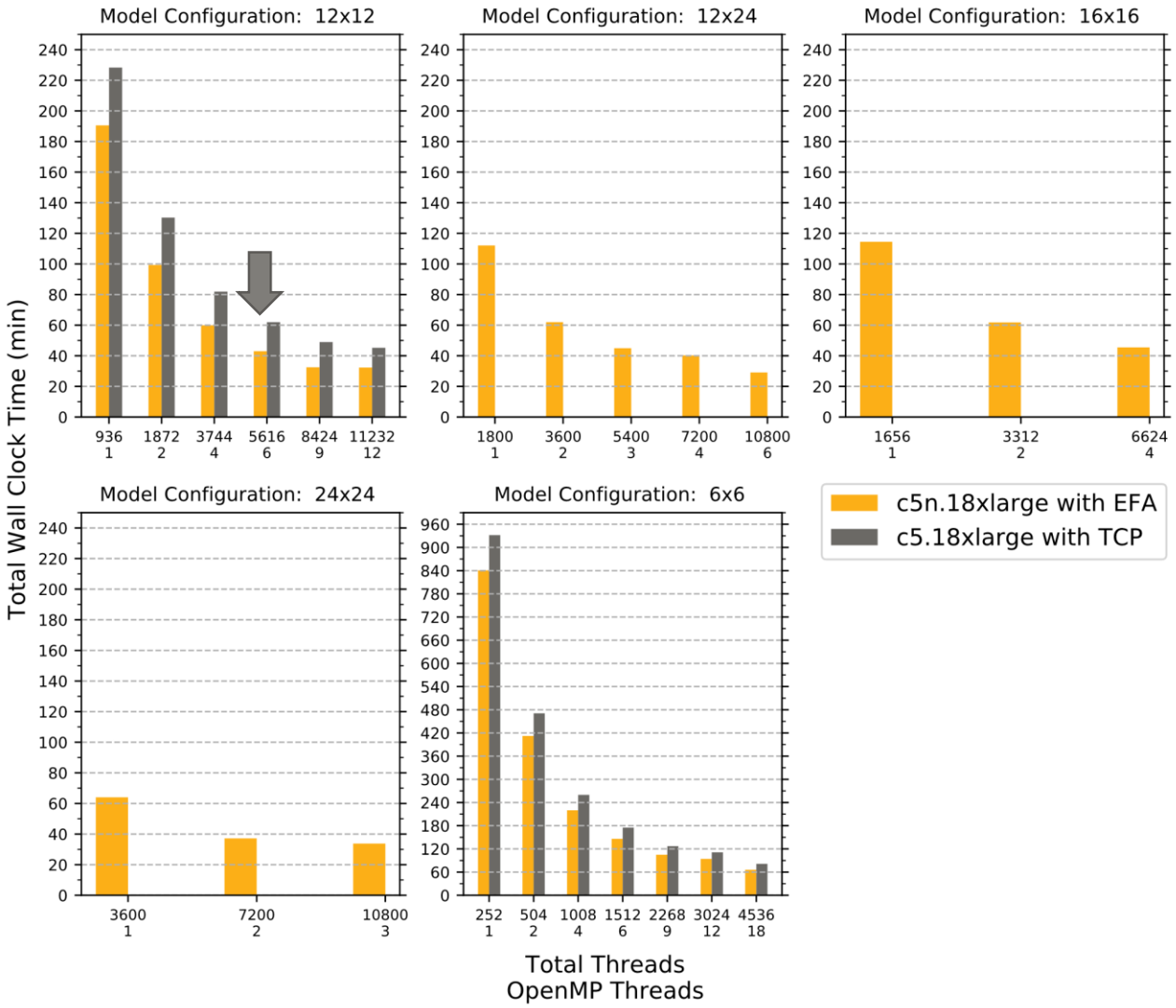
MAXAR

Results – FV3GFS total wall clock time

- Smaller model configurations (6x6) rely too heavily on OpenMP threading to boost total application thread count.
  - Performance lacks compared to other configurations.

- 16x16, albeit configurable to c5/c5n instances, has the worst performance of the larger configurations for varying core counts.

- The 12x24 configuration exhibits strong scalability at higher MPI thread counts with varying OpenMP shows potential for larger MPI threading compared to the production configuration (12x12).



Final Thoughts

Conclusions

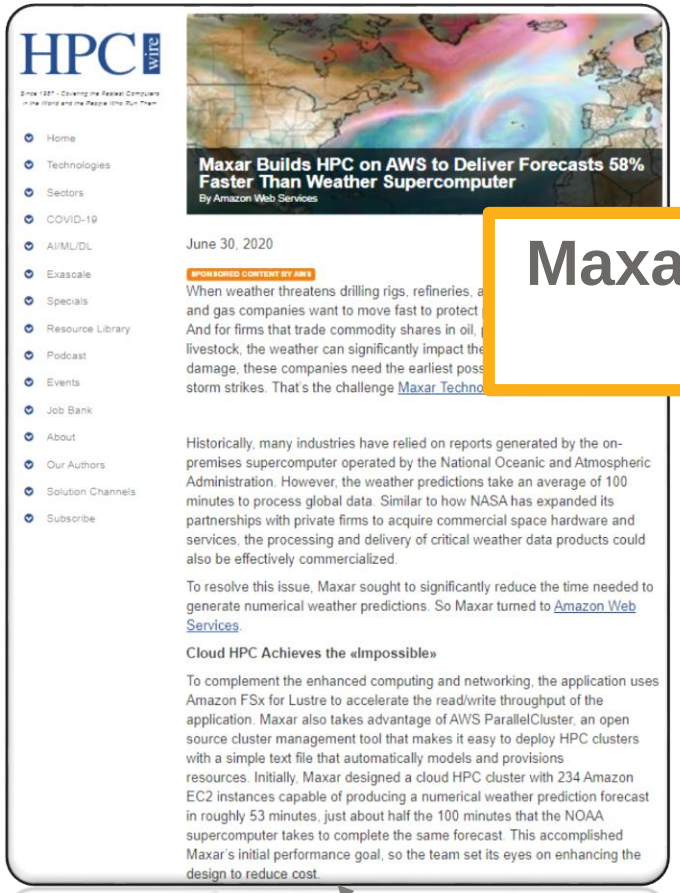
- The FV3GFS NWP application is scalable from 100s to 11,000+ cores, and potentially even more, on AWS Cloud HPC configured with c5n.18xlarge instances and EFA networking.
  - The scalability of the application is highly dependent on model configuration and the balance between MPI and OpenMP threading.

- Maxar, with its production system, cut the total run time of the FV3GFS by over **50%** compared to NOAA.

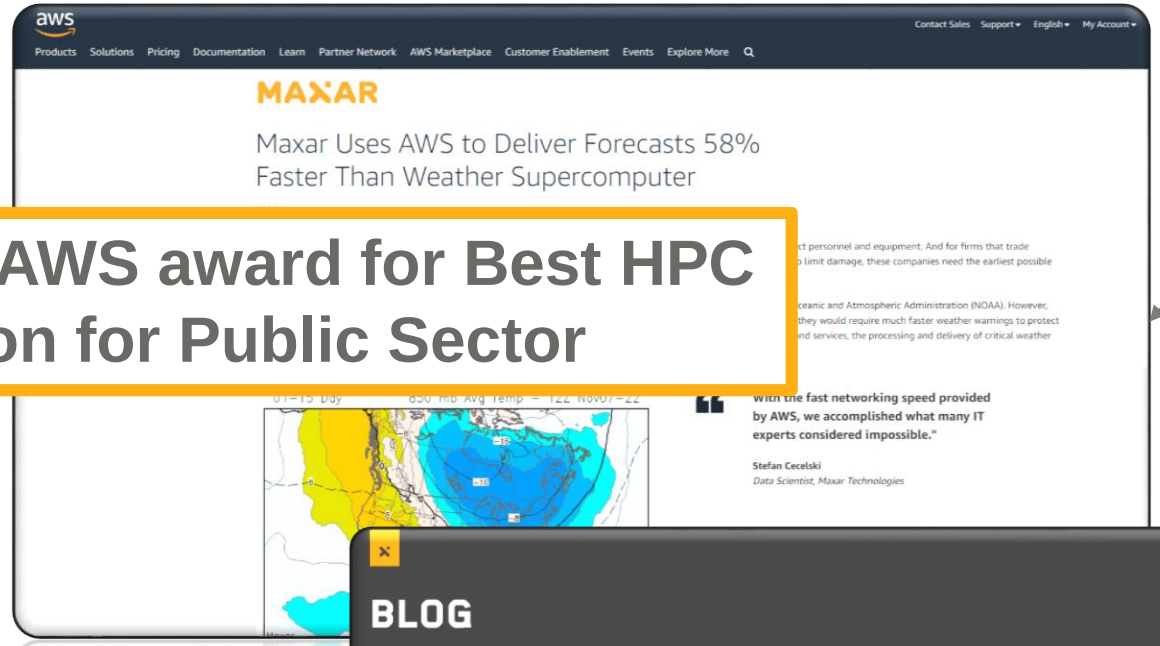
- Cost to develop and operate the production solution is significantly below the $500+ million, 10-year program for NOAA's bare metal supercomputer.

Awareness of Maxar's Cloud-Based HPC Solution for NWP

Maxar wins AWS award for Best HPC Solution for Public Sector



HPC Wire Article



AWS Published Case Study

Maxar Blog on Cloud Based HPC for NWP

MAXAR

MAXAR.COM