Lessons Learned About Running Weather, Climate, and Air Quality Models on the Amazon Cloud

Matthew J. Alvarado, Ethan H. Fahy, George Modica, Erik J. Fanny, Chantelle R. Lonsdale, Elizabeth S. Bettencourt, Rebecca Adams-Selin, Thomas Nehrkorn, and Jon Holdzkom

Atmospheric and Environmental Research (AER)

NCAR 2017 Modeling Research in the Cloud Workshop
Boulder, Colorado
May 31, 2017

malvarad@aer.com

“Cloud First” Strategy

- AER’s parent company, Verisk Analytics, has decided on a “cloud first” corporate IT strategy.
  – Shifts capital costs to operational costs.
  – Outsources hardware maintenance.
  – Can easily scale up and down with demand.
  – You don’t have to pay for idle resources.

Compute
- EC2
- EC2 Container Service
- Lightsail
- Elastic Beanstalk
- Lambda
- Batch

Storage
- S3
- EFS
- Glacier
- Storage Gateway

Database
- RDS
- DynamoDB
- ElastiCache
- Redshift

Networking & Content Delivery
- VPC
- CloudFront
- Direct Connect
- Route 53

Migration
- Application Discovery Service
- DMS
- Server Migration
- Snowball

Developer Tools
- CodeCommit
- CodeBuild
- CodeDeploy
- CodePipeline

Management Tools
- CloudWatch
- CloudFormation
- CloudTrail
- Config
- OpsWorks
- Service Catalog
- Trusted Advisor
- Managed Services

Security, Identity & Compliance
- IAM
- Inspector
- Certificate Manager
- Directory Service
- WAF & Shield
- Compliance Reports

Analytics
- Athena
- EMR
- CloudSearch
- Elasticsearch Service
- Kinesis
- Data Pipeline
- QuickSight

Artificial Intelligence
- Lex
- Polly
- Rekognition
- Machine Learning

Internet Of Things
- AWS IoT

Game Development
- Amazon GameLift

Mobile Services
- Mobile Hub
- Cognito
- Device Farm
- Mobile Analytics
- Pinpoint

Application Services
- Step Functions
- SWF
- API Gateway
- Elastic Transcoder

Messaging
- SQS
- SNS
- SES

Business Productivity
- WorkDocs
- WorkMail
- Amazon Chime

Desktop & App Streaming
- WorkSpaces
- AppStream 2.0

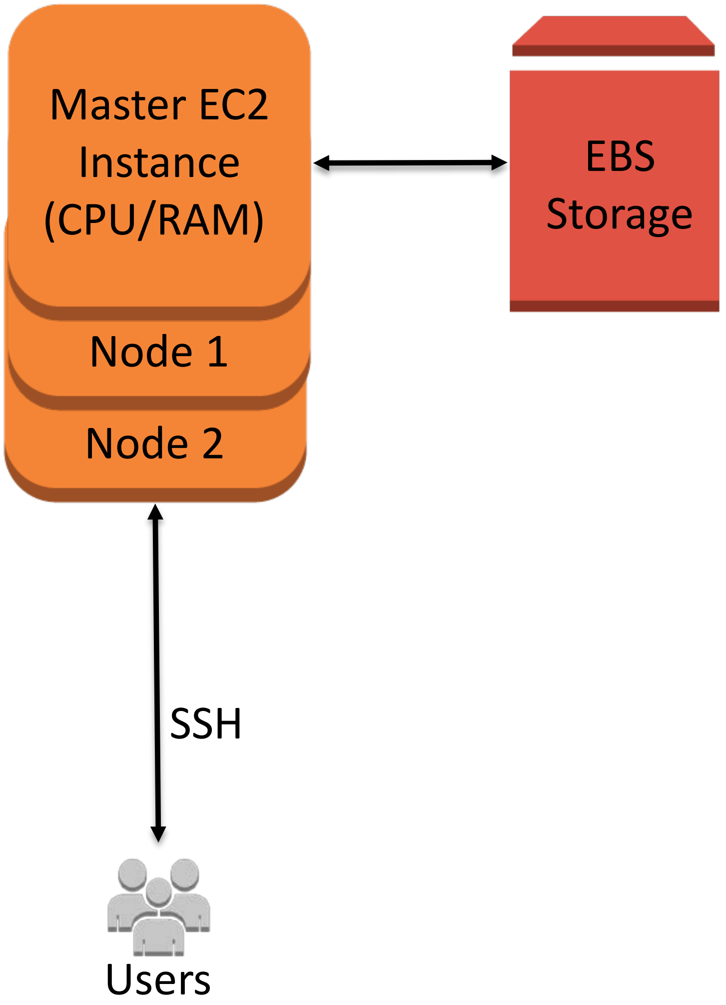
So you've been told to move to the cloud. Just log on to the AWS Management Console and unlock power beyond your wildest dreams!

OK Let's see, I'll just log on to the console and …

Atmospheric and Environmental Research

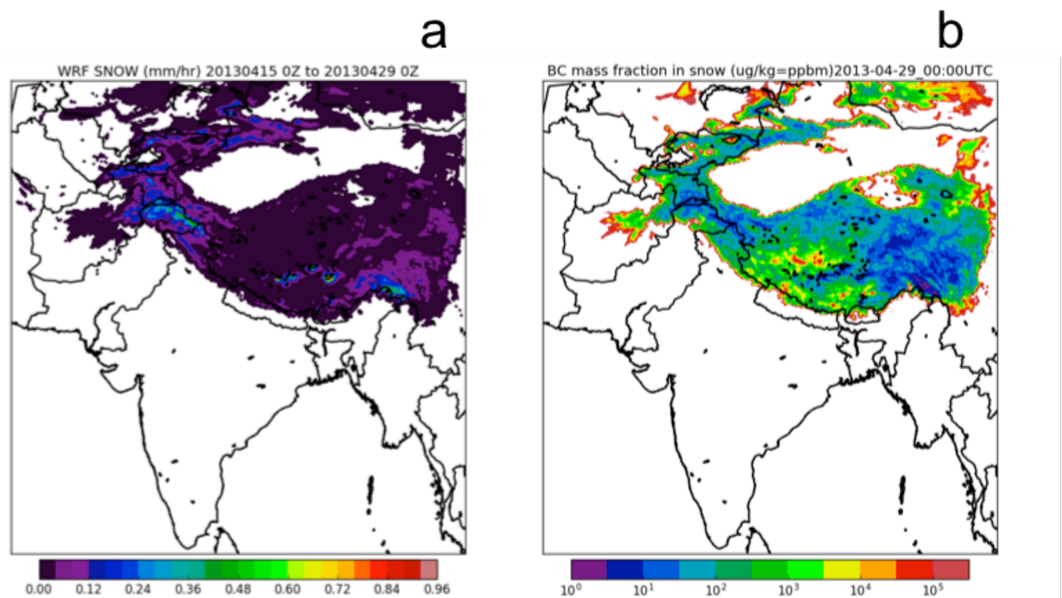
"Simple" Cloud Computing

- Replicate your current servers on the cloud.
- EC2 Instances
  - Equivalent to running a virtual machine in the cloud.
  - You can create an Amazon Machine Image (AMI) or use an existing AMIs as a starting point.
  - ***But you are responsible for configuring and maintaining the entire system after that.***
- cfnCluster or other tools can allow you to add and remove EC2 instances as cluster nodes.
- Can clone duplicates.
- Tags make it easy to track charges.



Example Project: Future black carbon deposition to the Himalayas

- Run WRF-Chem with GISS ModelE initial and boundary conditions
  - 12 km resolution and 30 vertical levels, tagged BC tracers
  - 4 months per year, 3 years (ENSO phases)
  - 2 emission scenarios
  - 24 month-long runs
- 6 month period of performance
- Would not have been possible on AER internal computer resources.



Example from WRF-Chem run on the cloud. (a) Average snowfall (mm/hr). (b) Mass ratio of BC in fresh snow (µg BC/kg snow).

Challenge #1: Choosing instances

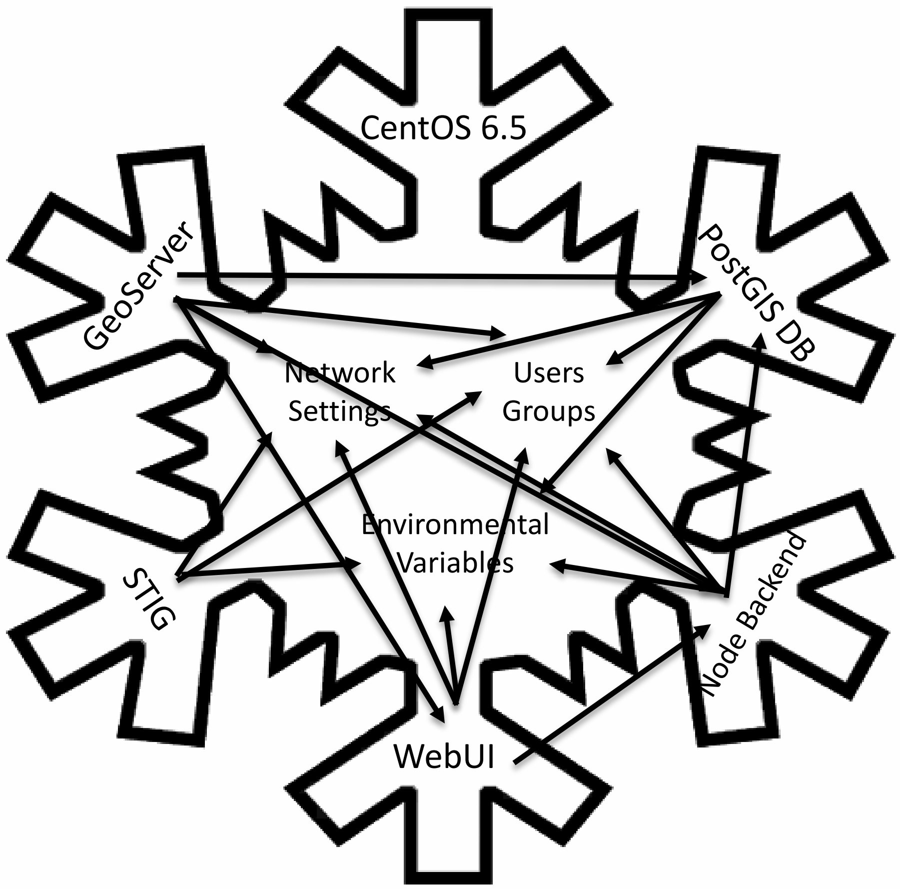
- **The cloud gives you a lot of CPUs, but that may not be your problem.**
  - It turned out that we were highly memory-limited (lots of segmentation faults on small instances).
  - Therefore we had to use 5 c4.8xlarge instances for each model run to get sufficient RAM.
  - Now you can use new memory- or storage-optimized instances if you are memory or I/O limited.
- **Price is rounded up to the nearest hour!**
  - If you find a bug and can fix it quickly, you might be better off leaving the instances attached.

Challenge #2: Maintaining the server

We need to update the version of GeoServer

We need to add a new user

HEARTBLEED!?! ?!?!?!?! ALSO FLASH!?!?!?!?!



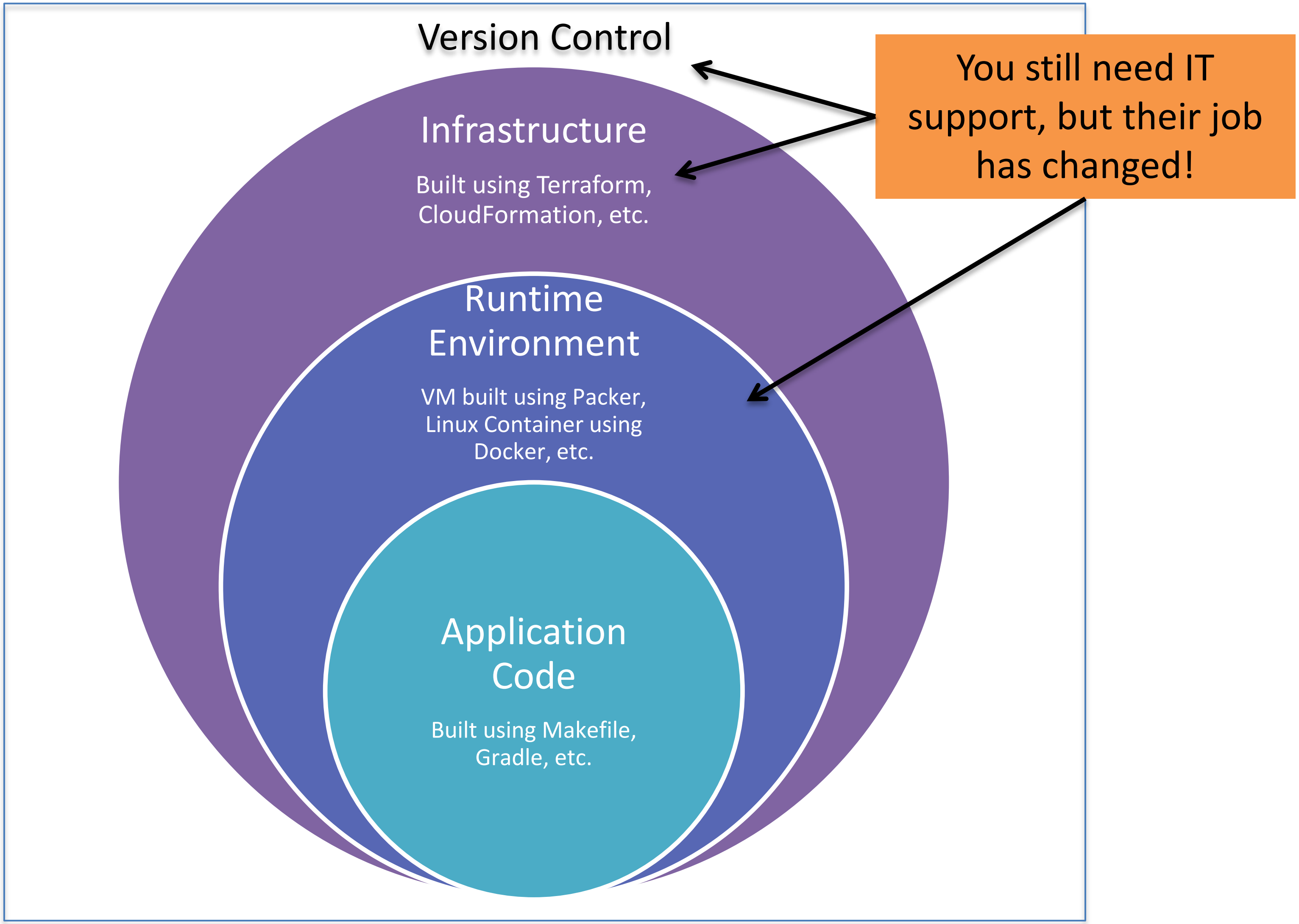
This hard drive sounds like a cement truck today…

Can we upgrade this to CentOS 7 by tomorrow?

Can we load balance this whole thing?

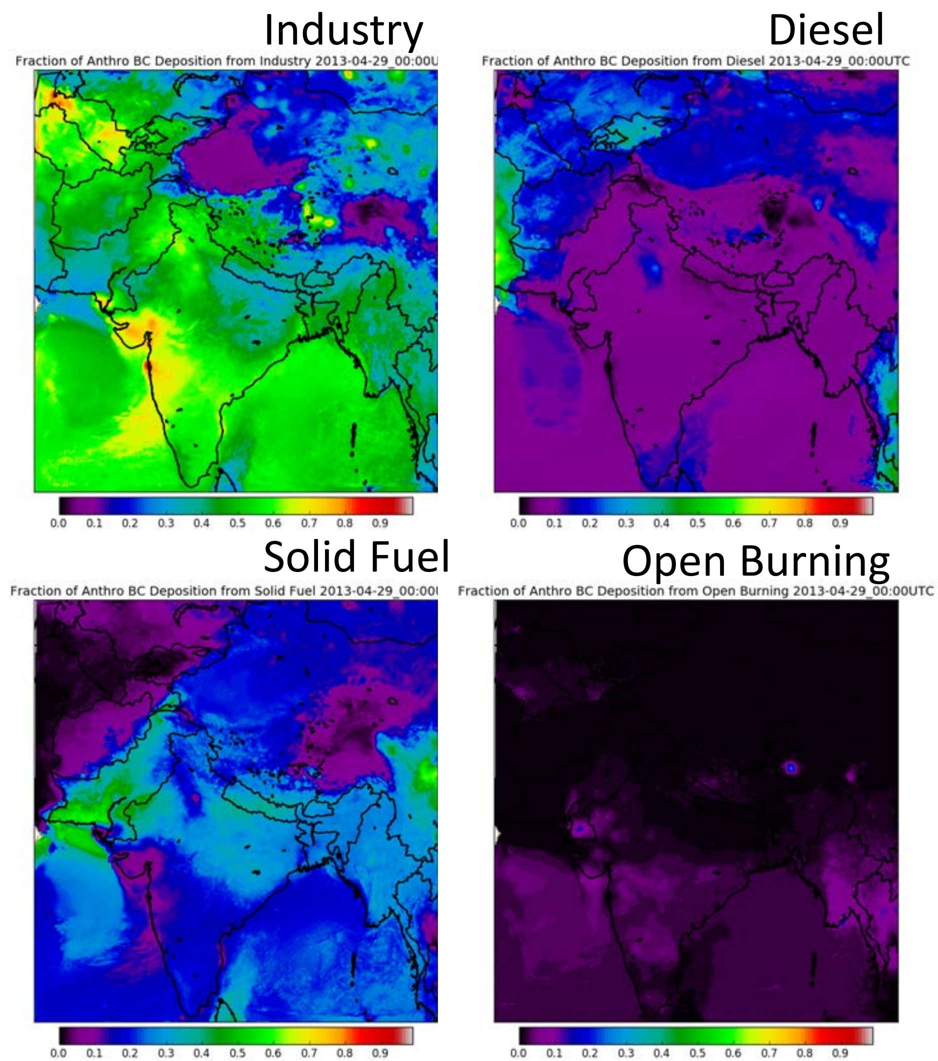
Atmospheric and Environmental Research

Solution: Infrastructure as Code



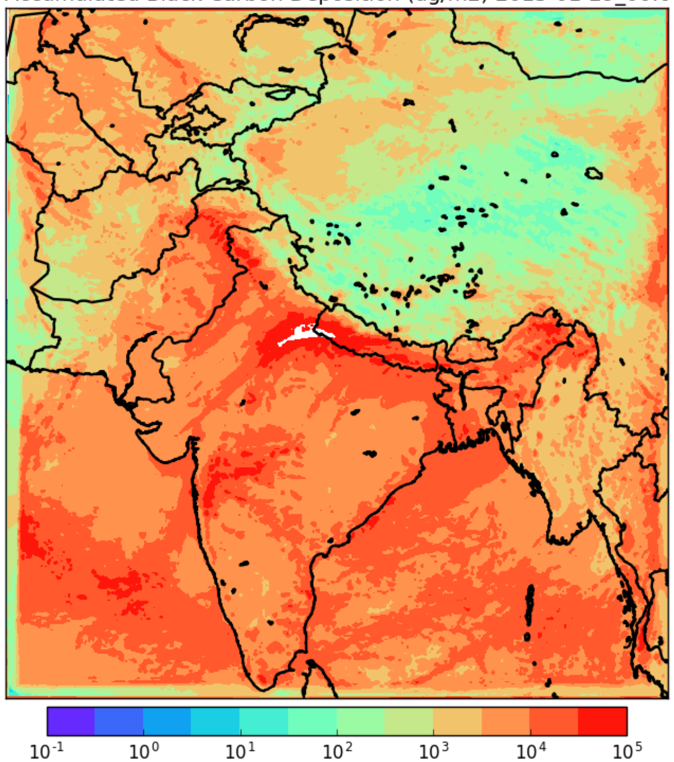
Challenge #3: Data Storage

- At least 658 GB needed for each model month run
  - 70 GB: wrfinput and wrfbdy files
  - 336 GB: wrfout files every 6 hr
  - 252 GB: wrfrst files every day
- So we need ~**15.8 TB!**
  - EBS: $100/TB/mo.
  - S3: $22.50/TB/mo.
  - Glacier: $4/TB/mo.
- Our approach:
  - Limit EBS storage to 4 parallel runs.
  - Move data from EBS to S3 as quickly as possible.
  - Don't keep every restart file.
  - Dump data to Glacier at end of project.
  - Note: Data transfer between EBS volumes and between EBS and S3 can be slow enough to be rate-limiting.



Challenge #4: Data Retrieval

Total Accumulated Black Carbon Deposition (ug/m2) 2013-01-29_00:00UTC



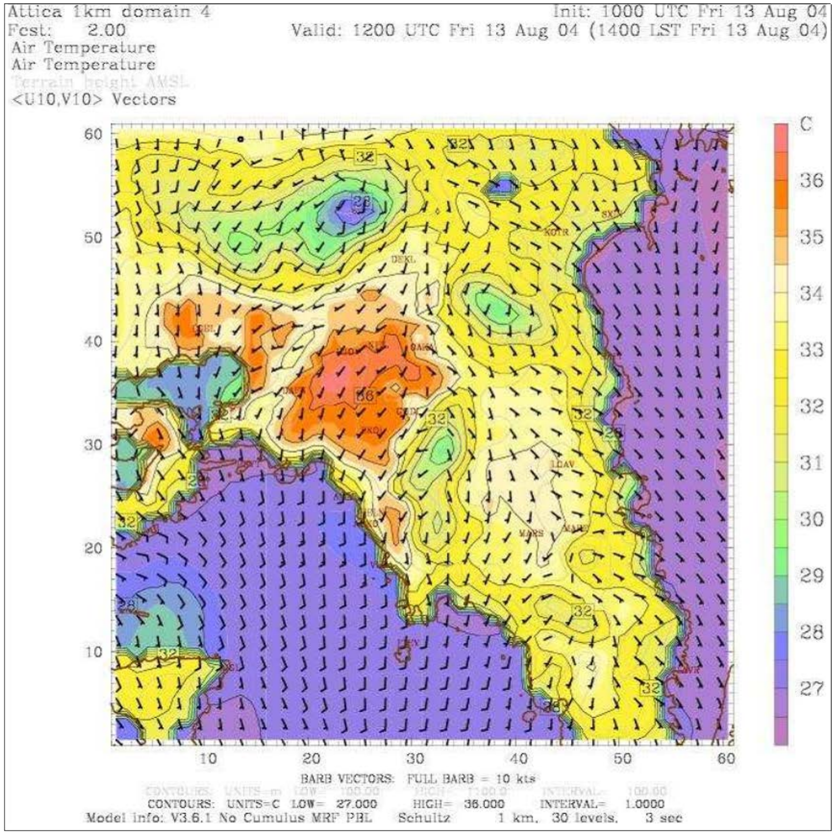
- AWS downloads cost $90/TB.
- Do post-processing and analysis on the cloud.
  - Use Python and R to avoid license charges
  - Use THREDDS to make data accessible via netCDF viewers like Panoply or via the web.
  - Don't forget to add the packages to your version-controlled infrastructure!
- Need to sort data between:
  - What you need to bring home.
  - What you want to keep on the cloud.
  - What you are willing to delete and regenerate.
- **None of these choices are free, and which is best depends on the application!**

More Advanced Cloud Computing: nCast for WRF and AQcast for WRF-CMAQ

- ✓ Scalable systems on AWS Cloud
  - Run arbitrary number of model runs on identical machines.
  - Configuration (OS, compilers, etc.) saved as Terraform modules.
- ✓ Automated systems
  - Gathering and preparation of model inputs and post-processing of model outputs done automatically.
- ✓ Unified systems
  - All components run on a common architecture using common XML input specification file.
  - Removes chances for inconsistencies in specifications.

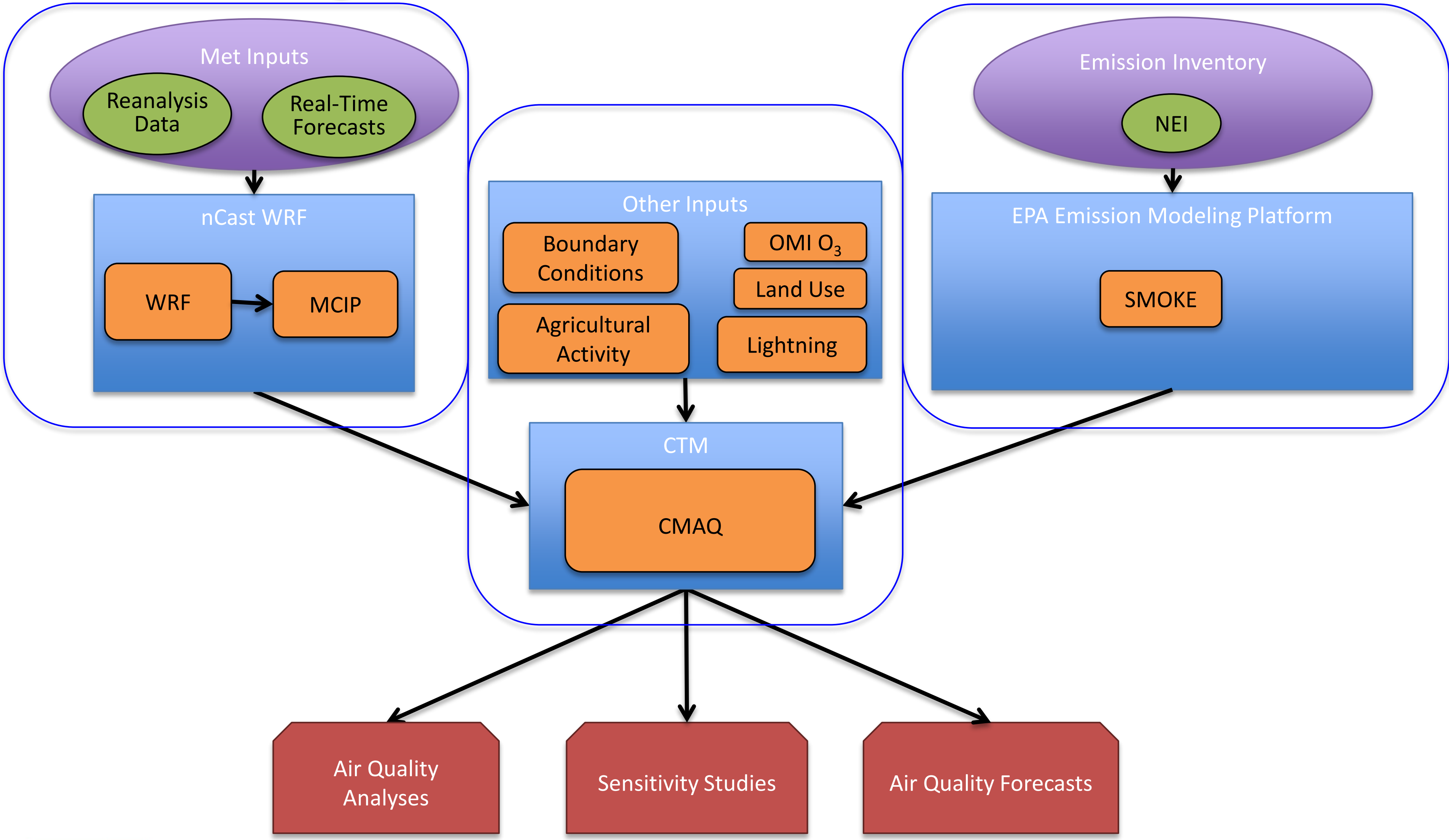
nCast WRF Forecasts for Greek Olympics

Fig. 5: *Example of nowcasting system graphic showing forecast of 2-m temperature (shaded, degrees C) and 10-m wind field (using standard notation of one barb=5ms^{-1}).*



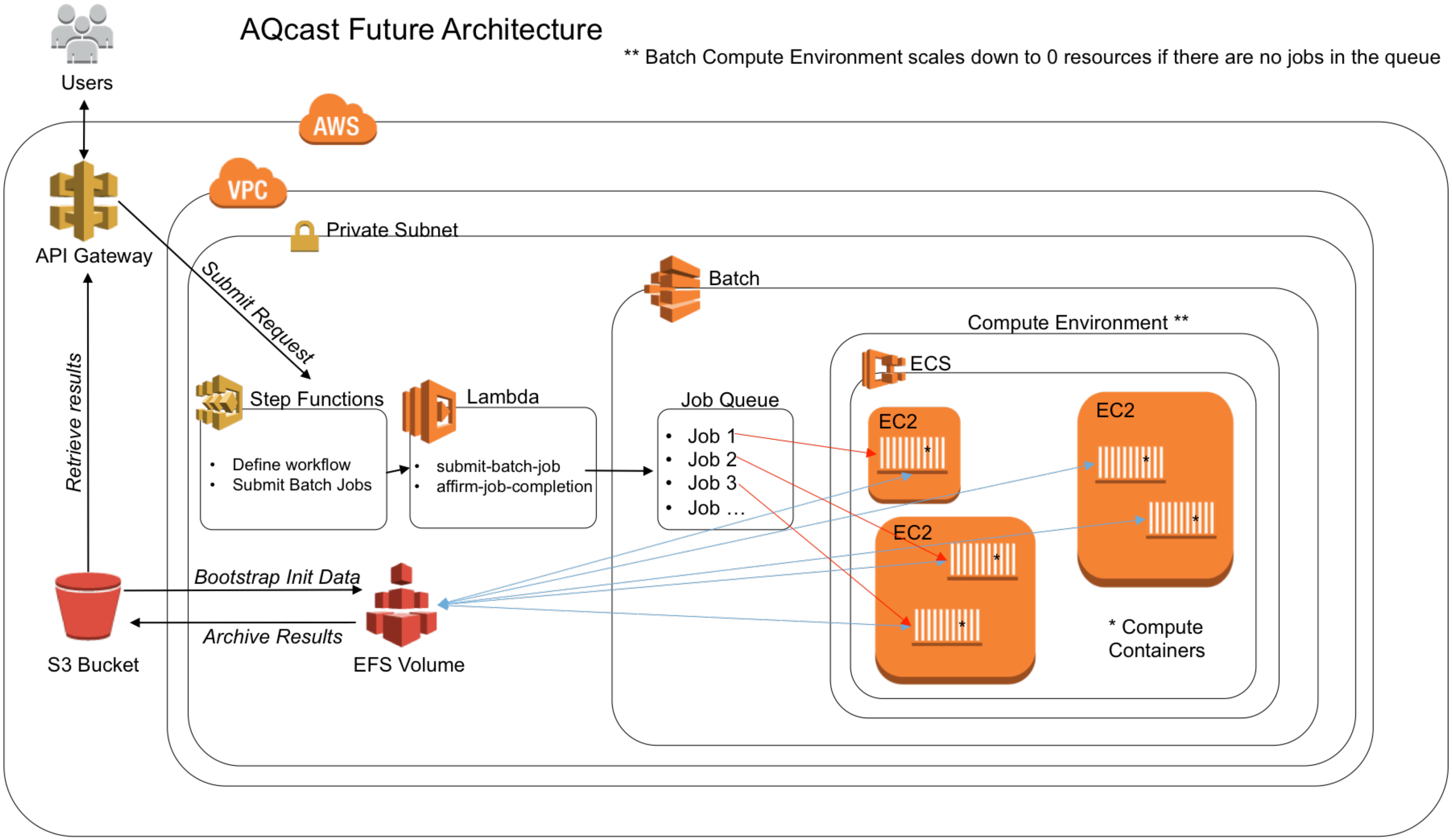
aer Atmospheric and Environmental Research

AQcast System for Continental US



Atmospheric and Environmental Research

Amazon Batch allows you to automate instances and storage transfers



aer
Atmospheric and Environmental Research

Summary: Lessons Learned

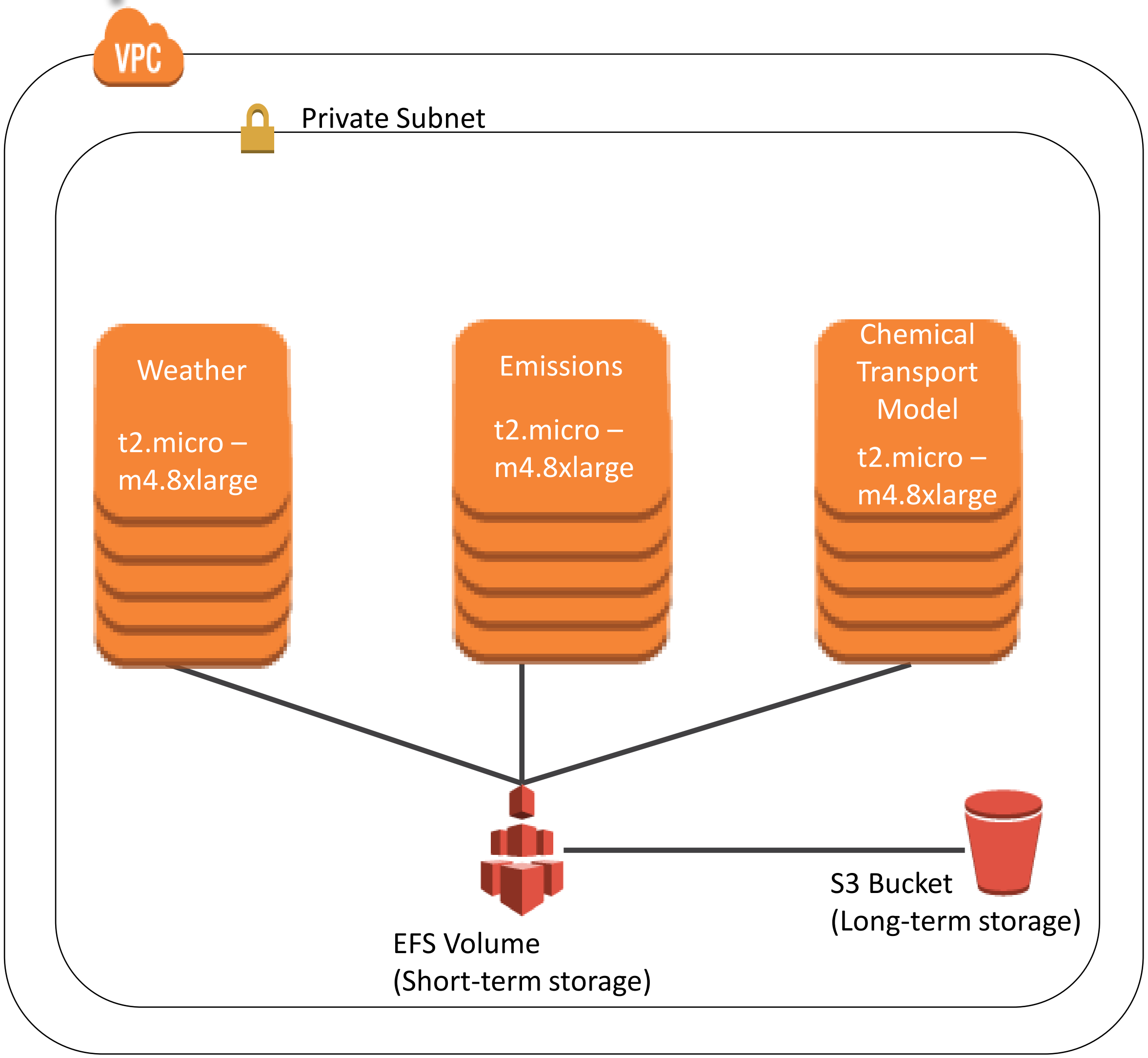
- Cloud computing allows you to quickly run an arbitrarily large number of model simulations.
- Allows you to share your whole system with other users – don't need to worry about different compilers or OS!
- But it also allows you to blow through your entire project budget in an afternoon, so you need to plan carefully!
  - – Save your infrastructure as version-controlled software.
  - – Choose the correct instance for your job.
  - – Examine if it is cheaper to regenerate or store the output.
  - – Your computers aren't "free" anymore, so get as precise a cost-estimate as you can for your proposal/project, and then add a buffer.
    - AWS Cloud Calculator
    - NCAR Performance statistics for CESM, etc.

What support do scientists need to use cloud computing?

- Many of the cloud infrastructure tasks are better handled by IT specialists rather than scientists.
- But the IT specialists must understand the models you are using, so that they can help choose reasonable compute instances and storage types, automate routine tasks, etc.: ***"DevOps" practices***
- At AER we have a team of software engineers and IT experts who regularly work with atmospheric, ocean, and climate scientists and are able to provide the above support.
- We are exploring if there is a market for us to provide others with our own cloud-based systems for running WRF & CESM: ***Infrastructure as a Service***
- Interested? Email malvarad@aer.com

BACKUP SLIDES

Implementation on Amazon Web Services



- Components can be run separately or together.
- Preexisting input met or emission files used by uploading them into the EFS volume.
- Currently, starting and stopping machines and the transfer the files between storage volumes is done by the user, but we are working to automate that as well.

EC2 Instances

- Equivalent to running a virtual machine in the cloud.
- You can create an Amazon Machine Image (AMI) or use one of many existing AMIs as a starting point.
- ***But you are responsible for configuring and maintaining the entire system after that.***

Instance	CPUs	RAM (GB)	Price (per hr)
c4.large	2	3.75	$0.10
c4.xlarge	4	7.5	$0.20
c4.2xlarge	8	15	$0.40
c4.4xlarge	16	30	$0.80
c4.8xlarge	36	60	$1.59

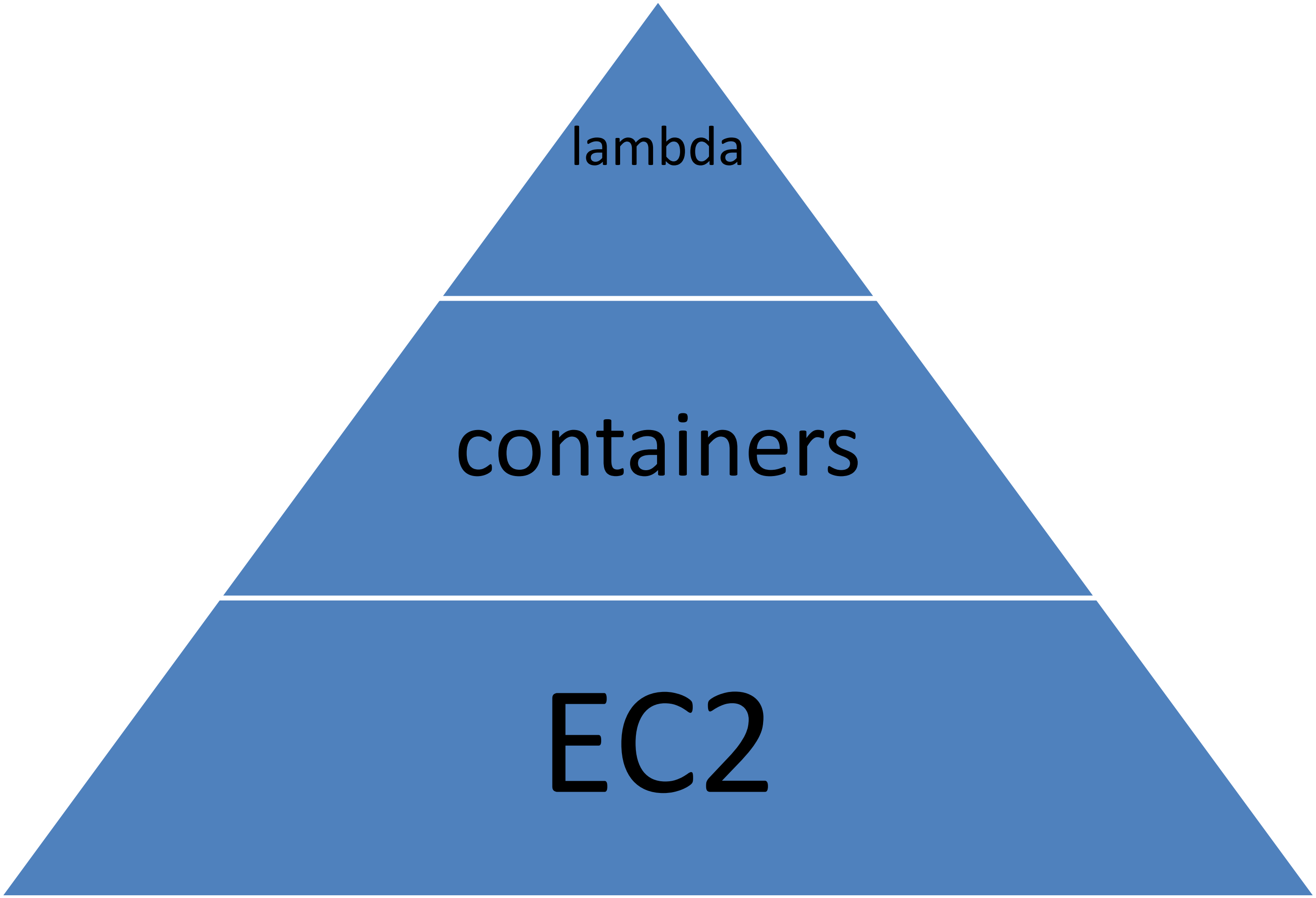
AWS Data Storage

RDS $$$$	•Managed relational database •You pay for underlying EC2 and EBS resources + I/O •Postgres, MySQL, Oracle, MariaDB, Microsoft SQL Server
EFS $$$$	•NFS 4.1, mounts to EC2, etc. •Connect to many machines at once •Expands and contracts from 0 to 8 exabytes (!)
EBS $$$	•Analogous to a local hard drive that can be mounted to EC2 •Can only be mounted to one EC2 instance at a time •Must preallocate storage space
S3 $$	•Object store, cannot be mounted and used like a regular block store.
Glacier $	•Like S3 but even cheaper •Used for cold storage of files that do not need to be regularly accessed.

$ not drawn to scale

Atmospheric and Environmental Research

AWS Compute Primitives



EC2

- Equivalent to running a virtual machine in the cloud.
- You are responsible for configuring and maintaining the entire operating system.
- Can be autoscaled.
- You can create an AMI or use one of many existing AMIs as a starting point.
- https://aws.amazon.com/ec2/pricing/on-demand/

Linux Containers

- Linux containers (usually Docker containers) are like having lightweight sandboxes that can run within an operating system.
- Define/build Docker Image using Dockerfile.
- Share Docker images through a registry such as ECR or Docker Hub.
- Shared images will work on anyone’s machine; no installation instructions or dependency management to worry about.
- Containers are the computational unit for AWS Batch.
- Containers can be deployed in AWS EC2 Container Service (among other places).

AWS Lambda

- You define a function and upload it to AWS.
- You pay only when your function is executed.
- You can execute as many instances of this function as you want and never worry about autoscaling or managing anything.
- There is a 5 minute runtime limit per function execution.
- You don’t have a local development environment; lambda functions only run on AWS.
- Limited language support, but you can wrap most binaries with those limited languages.

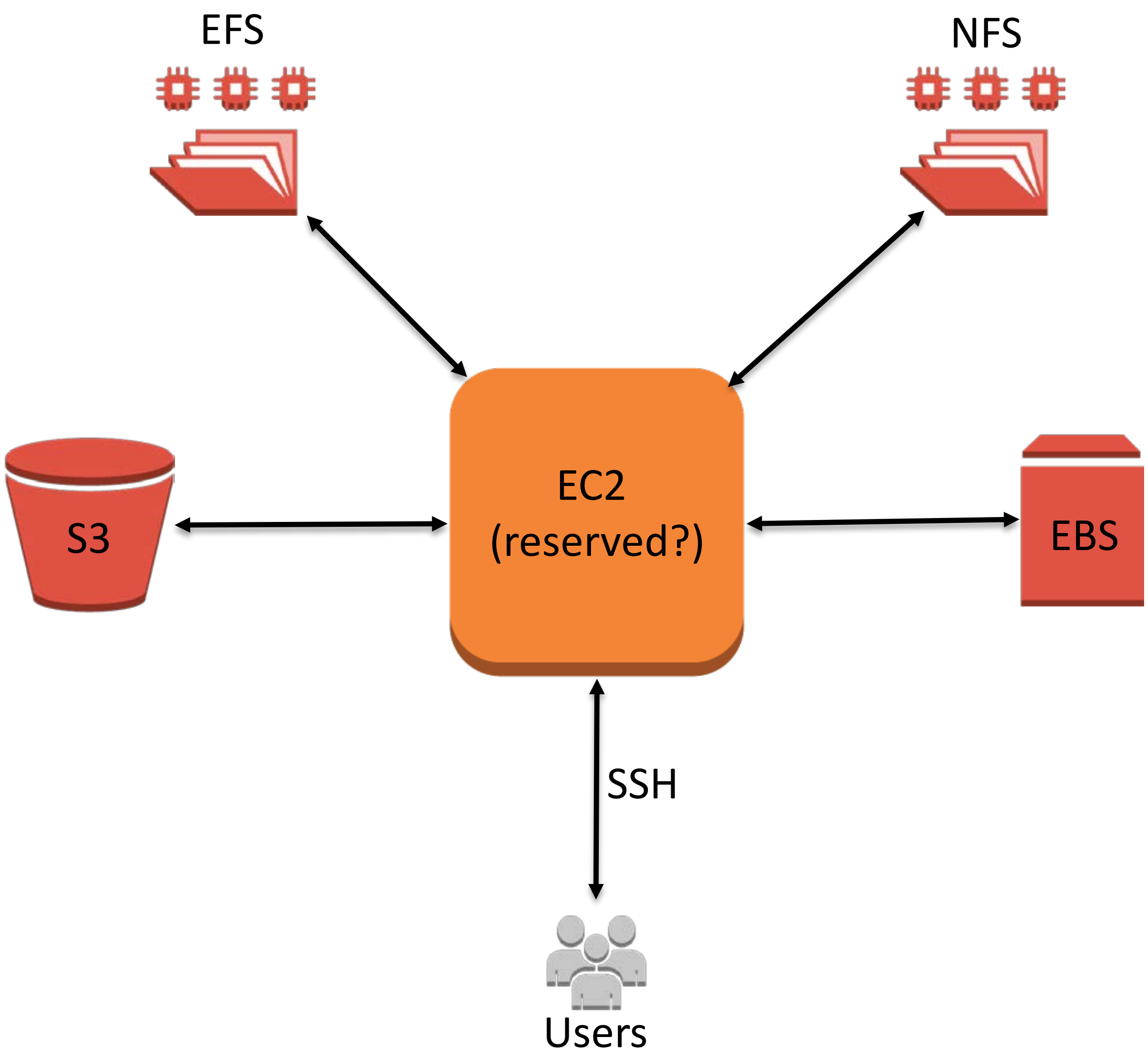
Code must be in a Container and able to run on a single instance to use Batch (for now!)

- Linux containers (usually Docker containers) are like having lightweight sandboxes that can run within an operating system.
- Shared containers will work on anyone’s machine; no installation instructions or dependency management to worry about.
- Containers can be deployed in AWS EC2 Container Service (among other places).

AWS Storage Gateway

- Cloud be used to store underlying data in S3 but provide a cached local copy of latest files as a mountable drive on the local network
- This could be useful for opening output data using local GUI-based tools.

"Rotor" on the Cloud



What if you have HPC jobs too big for any one EC2 instance?

- AWS Batch is currently limited to running jobs that fit on a single EC2 instance; if you need more than that, you might consider:
  - cfnCluster
  - StarCluster
  - Cycle Computing

Proposed Solution: Two levels of resources

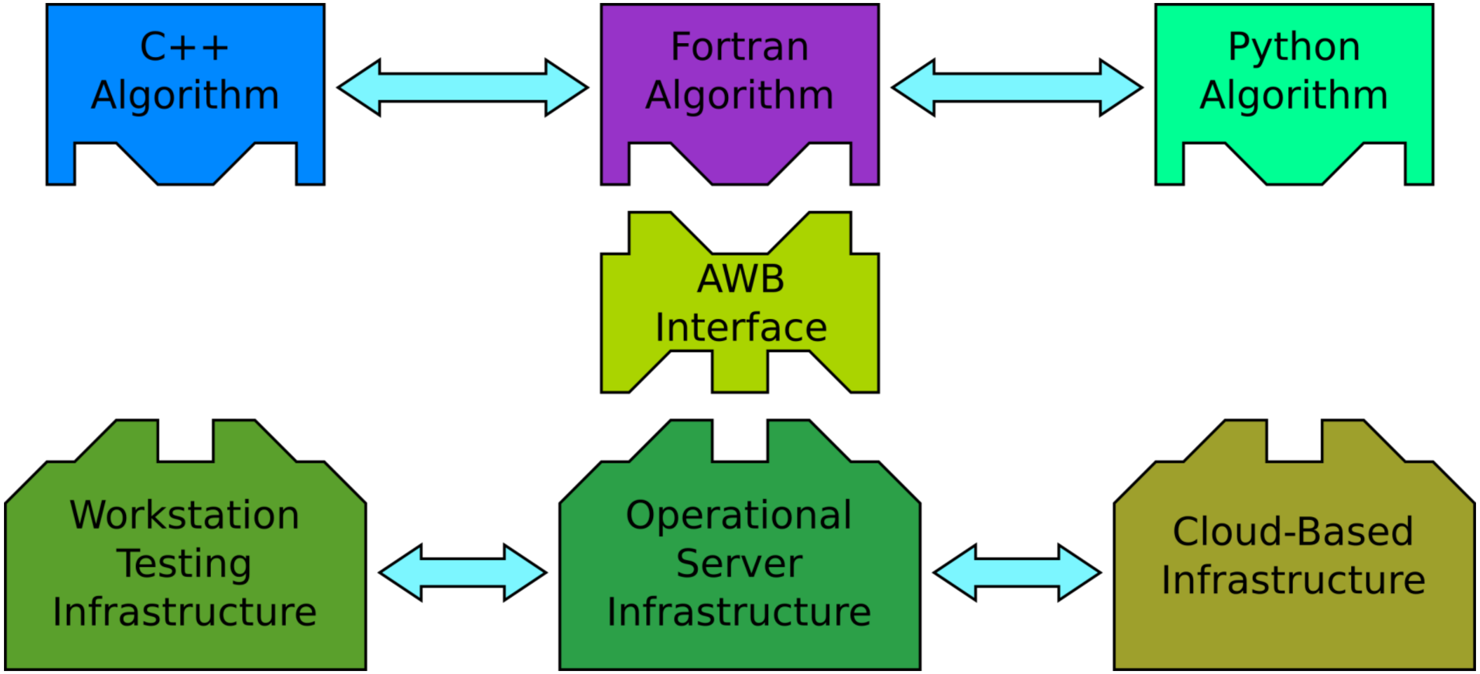
Community Resources

- Freely available, paid for by general “tax” on projects.
- Number and type of instances is fixed.
- Amount and type of storage is fixed.
- Have to share with your colleagues, wait your turn.
- Imposes a natural limit on costs, akin to limits of in-house servers.

Project-Specific Resources

- All charges are applied to your project.
- You can get as many instances and as much storage as your project can afford.
- Only used by your project.
- Have to do a cloud spending plan as part of your proposal. Use the community resources to perform small tests to help estimate the costs.

Future: Modular Algorithm Interfaces



- Initially designed for satellite data processing.
- Standardized interface allows algorithms and infrastructure to be swapped without code changes.
- Algorithms in different languages run in the same environment.
- Users can change/add algorithms using standard interfaces.
- ***Could allow scientists to test new algorithms while using containers for the rest of the code?***