NERACOOS Infrastructure Update: A Steller view of problematic data

Alex Kerney Dylan Pugh



Gulf of Maine Research Institute

Science. Education. Community.

Where Are We?



NERAC

NORTHEASTERN REGIONAL

·....

OCIATION

Gulf of Maine

Where We Started

- Cron*
- SEC cron but less likely to start up correctly with the system
- Untracked, unmanaged, conflicting environments
- Observability = cranky emails
- Super rickety EC2 instances that were migrated from another provider

Gulf of Maine

Research Institute

*if you are still in the world of cron, or other opaque orchestration systems, <u>healthchecks.io</u> and similar services will help you sleep better.

What Were Our Goals?



- Know what state the data is in
- Know what are workflows were up to
- Allow services and data pipelines to evolve at their own rate
- Manage the blast radius, have a better idea of what failures will affect
- Have standard ways to describe, deploy, observe, and interact with parts of the system

Essential vs Accidental complexity (aka The Mess)

Essential

- Dealing with other systems
- Derived data
- Services have to run somewhere
- Services & data needs to be publicly accessible
- Services & data have dependencies
- The ocean

Accidental

 Lack of observability/monitoring/ debug-ability

Gulf of Maine

Research Institute

- Library dependency conflicts
 between workloads
- Un-declared data dependencies
- Network or server hiccups
- Cron timing mismatches
- Blurred lines between infrastructure and data complexity

Early Decisions - Kubernetes

- Docker containers enforce strong isolation and repeatable environments
- Kubernetes enforces strong interfaces between workloads
- Lots of community, tooling, and knowledge
- Standard ways to see:
 - Which workloads are exposed to the internet
 - The dependencies of a workload
 - States of components

kubectl get --all-namespaces ingress

NAMESPACE	NAME	CLASS	HOS
argocd	argocd-server-ingress	<none></none>	arg
buoy-barn	buoy-barn-ingress	<none></none>	buo
inundation-dashboard	inundation-tool	<none></none>	inu
kube-prometheus-stack	kube-prometheus-stack-grafana	nginx	gra
mariners-dashboard-dev	mariners-dashboard	<none></none>	mar
mariners-dashboard	mariners-dashboard	<none></none>	mar
mbon-erddap	mbon-erddap-ingress	<none></none>	ism
oa-erddap	oa-erddap-ingress	<none></none>	oca
sea-eagle	dagster-ingress	nginx	sea
sea-eagle	dagster-oauth2-proxy	nginx	sea
sea-eagle	data-erddap	nginx	dat
sea-eagle	data-nginx	nginx	dat

argocd.aws.neracoos.org buoybarn.neracoos.org grafana.aws.neracoos.org mariners-dev.aws.neracoos.org mariners.neracoos.org ismn.erddap.neracoos.org oca.data.neracoos.org sea-eagle.neracoos.org data.neracoos.org data.neracoos.org data.neracoos.org

ADDRESS

80, 443 a60188d8abf764f7599a9b430842dba3-1113587032.us-east-1.elb.amazonaws.com a60188d8abf764f7599a9b430842dba3-1113587032.us-east-1.elb.amazonaws.com 80, 443 80, 443 3d8ab†764†7599a9b430842dba3-1113587032.us-east-1 80, 443 a60188d8abf764f7599a9b430842dba3-1113587032.us-east-1.elb.amazonaws.com 80, 443 764f7599a9b430842dba3-1113587032.us-east-1 80, 443 a60188d8abf764f7599a9b430842dba3-1113587032.us-east-1.elb.amazonaws.com 80, 443 '64f7599a9b430842dba3-1113587032.us-east-1.elb.amazonaws.com 80, 443 a60188d8abf764f7599a9b430842dba3-1113587032.us-east-1.elb.amazonaws.com a60188d8abf764f7599a9b430842dba3-1113587032 us-east-1 elb amazonaws 80 443

PORTS

Gulf of Maine

Research Institute

Early Decisions - Kubernetes



- Spent 4 years learning on smaller projects and clusters as a team
- Learned how to lean on provider and community tools
 - Argo CD, Kustomize and Helm
 - Grafana, Cert-Manager, External Secrets Operator
- How do our services react to being containerized, treated as cattle, and forced into playing with the cloud?
 - ERDDAP & S3



What We Tried - Pachyderm

Language agnostic, Docker & Kubernetes based execution model

Gult of Maine Research Institute

- Workflows triggered by changes in data store structure
- Agnostic to type of data
- Very opinionated about structure of data store
- Really doesn't want anyone else touching its data store
- Minimal observability of tasks or data
- Licensing and cost went from friendly to extremely expensive

What We Tried- Prefect

922 1057 867 OF COASTAL

Gulf of Maine

Research Institute

- Python workflow engine
- Hybrid cloud options to let them manage the orchestrator
- Really quick to get going, and wrap your head around the concepts
- Easy to pass data between tasks
- UI is focused more on stats about your tasks and workflows, rather than what your data is doing
- Evolving towards more flexibility in task management than data management

¢	Summary	ூ 24 Hours 💌	Contract of the second	I9 upcoming runs
	In the last day 26 flow runs	96.2 % succeeded	✓ No reported failures in the last day Everything looks good!	UPCOMING LATE hourly-default-dataset-refresh > vermilion-whale • Scheduled for 1:00pm EDT • hourly-default-dataset-refresh > versatile-beetle • Scheduled for 2:00pm EDT • hourly-default-dataset-refresh > versatile-beetle • Scheduled for 3:00pm EDT • hourly-default-dataset-refresh > furry-walrus • Scheduled for 3:00pm EDT • hourly-default-dataset-refresh > wisteria-dinosa -

What We Tried- Dagster



- Python workflow engine (working towards supporting other languages natively)
- Hybrid cloud options to let them manage the orchestrator
- A little slower to get started with than Prefect
- Layered concepts, move between task or asset focused as needed
- Easy to pass data between tasks
- Easy to surface metadata to the orchestrator
- 'Data aware, but agnostic'

download_daily_json	daily_df Create a daily dataframe for Departme	 monthly_df Monthly dataframes aggregated from 	gy opstream data	datum_calculation_df		
All O A	112 0 0 ▲ 0 Materializing 1 partition b9eeb490	All O All All	Single dataframe for full timeseries for Materialized Sep 22, 11:50 AM	Format dataframe for analysis with NO Never materialized		
	尚 pandas	, pandas	i∥ pandas	© Upstream data	🗭 Upstream data	🕼 Upstream data
			🖬 station_info	🖬 standardized_df	🖬 ds_meta	Hohonu_tide_BoothHarbor_ME
			Download station info from Hohonu for	Standardized dataframe for full time s	Dataset metadata for Department of	ERDDAP data.neracoos.org/erddap/ c
				Materialized Sen 22 11:51 AM	Materialized Sep 22 11:51 AM	Materialized Sep 22 11:51 A

What We Didn't Try



- Argo Workflows Largely language agnostic, very little awareness of what is happening during a task
- Luigi Seems to have run out of steam, not very data aware
- Ploomer focused on orchestrating notebooks
- Kedro, Metaflow, Kubeflow focused on ML uses
- Airbyte, dbt, Snowflake
 - Often strongly data aware,
 - but 'n-dimensional data, what's that?' No awareness or ability to work with gridded data
 - Flyte also preferred tabular data

https://pixelastic.github.io/pokemonorbigdata/



Gulf of Maine

Research Institute

Where We Ended Up

NERACOO NORTHEASTERN REGIONAL ASSOCIAT OF COASTAL OCEAN OBSERVING SYSTE

Gulf of Maine Research Institute

Infrastructure - AWS

- EKS
- EFS
- S3
- System Manager -Parameter Store
- SQS & EventBridge

<u>Kubernetes</u>

- Argo CD
- Ingress-Nginx
- Cert-Manager
- External Secrets
 Operator
- Grafana
- Prometheus

Orchestrator

- Dagster
- Xarray
- Pandas
- Pydantic
- IOOS_qc
- Compliance
 checker

Declaring The Mess - Ops & Assets

Ops/Graphs/Jobs

- Ops (operations) are assembled into abstract graphs which can be reused
- Graphs can be customized into jobs
- Jobs can be started manually, or triggered by schedules or sensors

Assets

- 1 asset often equals 1 file
- Ops or graphs of ops under the hood
- Can be assembled into jobs or automatically materialized as dependencies change



Declaring The Mess - Partitions



...



- Most data has some sort of natural division
 - Calendar: Daily/Monthly
 - Deployment: Buoy M0133
- Mapping between partitions
- Materialization policies let Dagster know when to update a partition

••• > Hohonu / tide_gauges / boothba	ay_harbor_me / monthly_df 🗐 🖩 Asset in hohonu C 🖷
Partitions Events Plots Definition Lineage	Auto-materialize history
[2023-06-012023-09-01]	
2023-06-01	
4 Partitions Selected	
(A	2023-09-01 Materialized
2023-09-01	
2023-08-01	Latest materialization Run
2023-07-01	 Sep 22, 11:55 AM (2039 prior events) 871fbe49
2023-06-01	Metadata

```
daily partitions = DailyPartitionsDefinition(
    start_date=START_DT,
    end_offset=1,
monthly partitions = MonthlyPartitionsDefinition(
    start_date=START_DT.strftime("%Y-%m-01"),
    end offset=1.
Dasset(
    partitions def=daily partitions,
def daily_df() \rightarrow pd.DataFrame:
    """Transform daily response from Hohonu into a dataframe
     ...
@asset(
    partitions_def=monthly_partitions,
    ins={
        "daily df": AssetIn(
            partition_mapping=TimeWindowPartitionMapping(
                allow_nonexistent_upstream_partitions=True,
    auto_materialize_policy=AutoMaterializePolicy.eager().with_rules(
        AutoMaterializeRule.skip_on_not_all_parents_updated(),
def monthly df(
    daily_df: dict[str, pd.DataFrame],
  → pd.DataFrame:
    """Aggregate daily dataframes to a monthly dataframe""
```

Declaring The Mess -Concurrency & Retries



- Network connections and servers can be fickle
- If there is a hiccup, we want to retry, but not bring down the server
- Limiting simultaneous external connections makes other server admins happier

```
aop(
   retry policy=RetryPolicy(
        max_retries=3,
        delay=120,
        backoff=Backoff.EXPONENTIAL,
        jitter=Jitter.PLUS_MINUS,
def download file():
    """Download a file from the FTP server into the scratch directory""
    ...
ajob(
   tags={"one-off": "bio ww3 72",},
def bio ww3 72():
   Download Wave Watch 3 forecast from Bedford Institute of Oceanography
    and split up by regions
    ....
   downloaded = download_file()
```

Declaring The Mess -Schedules vs Sensors

Schedules

- Cron, but you know what's going on
- Predictable timing of sources
- Often for sources that you can ask for a specific subset

Sensors

- Waiting for events
- Unpredictable source timing
- Idempotent run_keys

	C			DOS LASSOCIATION RVING SYSTEMS		Gulf of Researc	Maine h Insti	tute
daily_hohonu_b	oothbay_harbor_n	ne 🚺 🛇 Schedule in I	hohonu C	🗿 Next tick: Sep 22, 6:	48 PM UTC ⊙ id: 481	22258	0:10 C Tes	st Schedule
Description	Update data for De	partment of Marine Reso	urces, Bootht	bay Harbor, ME Hoho	onu tide gauge following	g a cron schedule	(*/6 * * * *)	
Latest tick	Sep 22, 6:42 PM U	TC ③ 1 Requested						
Job	scheduled_hohonu	_boothbay_harbor_me						
Partition set	None							
Schedule	Every 6 minutes	*/6 * * * *)						
Execution timezone	UTC							
Tick history Run h	istory				Started	Requested	Failed [Skipped
Timestamp	Status	Runs			Requests			
bio_ww3_72_ft	p_sensor 💽 🕬	Sensor in bio_ww3@bio-w	w3 C Ō	Next tick: Sep 22, 2:51	PM		0:07 C	Test Sensor
Target bio	_ww3_72							
Cursor Nor	e 🖌 Edit							
Frequency ~5	min							
Recent ticks								
					•			
2:43 p.m.	2:4	4 p.m.	2:45 p.m.	2:46	p.m.	2:47 p.m.	2 Failed	548 p.m.
Timestame	Status	Curror		Pupe	-	Paquestr		
Sep 21, 7:53:54 PM	1 Requested	-		23eb0e67		nequests		
Sep 21, 4:37:26 AM	1 Requested	-		● e5721e0c				

Declaring The Mess -Dataset metadata

- A common structure for
 - Dataset services: ERDDAP/THREDDS/ Xpublish/STAC
 - QC tests
 - NetCDF attribute modification
 - Links to sources
- Allows reusing ops & assets across datasets
- Designed to be the source of data for a CKAN catalog

Gulf of Maine Research Institute title: NOS CO-OPS Currents Data, Cape Cod Canal, West End, 6-Minute services: erddap: - server: data.neracoos.org/erddap/ template_file: ../erddap.xml.jinja dataset id: PORTS CC realtime dataset_id_template: "PORTS_CC_realtime" type: tabledap thredds: - server: data.neracoos.org/thredds/ template_file: ../thredds.xml.jinja dataset_id: PORTS_CC_realtime dataset_id_template: "PORTS_CC_realtime" qc: gartod: contexts: - streams: CS: gartod: aggregate: null gross range test: suspect span: - 10 - 220

> fail_span: - 0.5 - 249 spike_test:

aggregate: null gross range test:

CD:

gartod:

suspect_threshold: 3
fail_threshold: 4
flat line test:

suspect_threshold: 0.05
fail threshold: 0.1

Reveal The Mess -Asset & Job state

- Easy to see the overall state of jobs, and lots of info available if digging in further
- Assets surface even more info that jobs quickly









🕥 Overview Runs A	ssets D	aployment						(Search.	8		\$
o_ww3_72@bio-ww3	2 •	Over	view									
_ww3_72	80											
_ww3_cleanup	0	Activit	y Jobs Schedules Sensors (lesources Backfills							1	0:14 0
_ports@cc-ports	1 +	Time	line Assets 🗅 10 of 10 🗸	Q. Filter by job name			1hr	6hr 12hr	24hr	*	Now	•
com@fvcom	6 •	lobs				9/22/2923						
nerate_servpo@config-gen	3 .	0000		29		New		3PM				
		C ec.,	ports									н +
honu	8 *	C ger	nerate_service_configuration_repo		_						2 3	11 -
		Y Ger	nerate_data_neraorg_ERODAP_config	• 1 2								
heduled_hohonu_bath_me	Ø	Y Ger	nerate_data_nerarg_THREDDS_catal:	9								
eduled_hohonu_boothbay_ha	r O	Ca hoh	honu		_						60	1 .
eduled_hohonu_gmri_pier_po	o O	E Ad	hoc materializations		16 4							
neduled_hohonu_st_george_m	e O	¥ sch	weduled_hohonu_gmri_pier_pertland_m	•								
roups		Y sch	eduled_hohonu_boothbay_harbor_me									
h_me		" sch	eduled hohonu st.george me					0				
othbay_harbor_me		🗅 ons	set_tide_gauges									3 +
rl_pler_portland_me		Ca um	aine_buoys								10	1 -
george_me		¥ 401	Lsbe37_002m									
on@mbon-datasets	4 +	10A 1	1_doppler_currents									
		101 T	1_sbe37_020m									
bc_platformsndbc-platform	6 1 1	% A01	Lsbe37,003m									
set_tide_gauet-tide-gauges	5.1	¥ 401	1_sbe37_010m									
seden 19 men and 1 menter		Y ADI	1_met									
iaine_buoys@umaine-buoys	11 •	9 AD1	sbe37.001m									
1 danalar currente	200	8.P 401	Lunning									
1 0 0		1885	Hide not started steps 2465 3885	0 : 55 : 06 0	: 87: 00 9:0	8-14 0.05 Prepa	ring (0)	E P	a	le-execu	te all (ŋ -
		-					Core 100					
-		-				Deca	trag tot					203
	-	_	milli_file_exte				a 640	crig.				1.2
					-	Errore	d (0)					
						No ati	pe save and	red				
(3 /2 Ty	/pe a step	subset (exc	ioad_dateset_config+*)	Hide unselected s	teps							
(D) (Filter		8	Y Levels (5) ~							0	創 Cop	y URL
MP OP	EVENT	TVPS	INTO.									
705 generate_dataset_ser	340		Wrote EXCOAP dataset stub to /ent/	fs/dataseta/UNai.ne/liu/ys/	AD1/Net/Ad1_Net.erddd	ap data_meracoos_org_arddap.a	e1					
vices												
vices	ASSET_PO	HALLINT.	Haterialized value data.neracoos.o	rg/erddap dataset configs	AB1_Met.							
			 Anti-Arfa-Adatasets/UMainer/B Anti-Arfa-Adatasets/UMainer/B Config Gataset type://DDFaale/f Config Gataset type://DDFaale/f Config Gataset type://DTFaale/f Config Gatas	www surveyset.com/sgs / AB avar/ABI/met/ABI.Met.endda roanicD/Files" datatetID="A ABG/FeladEveryABinutes: Laats/DRains/Buoys/ABI/me artsIvas (1). v., met., et. mcinetadataInetas.	s.ast (<u>sine Allet</u>) <u>s.deta reracoos org (</u> M.met" active="true" t meRegex>	erodaext) (1)						
.245 generate_dataset_ser	STEP. OUT	PLT.	Vielded output "result" of type "A	ny*. (Type check passed).								

6fe-4e63-8c9d-65d8a9ecbb1b/generate_dataset_services

Reveal The Mess - Asset Checks, Plots, & Metadata

- We are already exposing asset & op metadata
- Currently working on more direct **QARTOD & Compliance Checker** integration

Events Plots Def	nition Li	neage Auto-materialize	mlw meter	s count	26917						
T Type (2) ~			mlw_meter:	s mean	1.5677	405728721625					
+: Sep 22, 3:33 PM			mlw_meter:	s std	0.9837	743733892538					
443d7c69	① csv.t	ail									
 ★: Sep 22, 3:27 PM ● a9885d8c 	25955	dt	navd88_feet	forecast_navd88_feet	navd88_meters	forecast_navd68_meters	mbhw_meters	mhw_meters	mtl_meters	msl_meters	mlw_meters
+: Sep 22, 3:21 PM	26967 26968	2023-09-22T23:42:00 2023-09-22T23:48:00	nan	0.55	nan nan	0.168	nan	nan nan	nan	nan	nan nan
+: Sep 22, 3:15 PM	26969 26970	2023-09-22T23:54:00 2023-09-23T00:00:00	nan nan	0.31 0.21	nan nan	0.094 0.064	nan nan	nan nan	nan nan	nan nan	nan nan
• 3f99a7fb					ST 552					Сору	Close



Check name	Status
attenuated_signal_test	⊘ Passed
climatology_test	Passed
flat_line_test	🛆 Warning
gross_range_test	🛆 Warning
location_test	Passed
qartod_rollup	⊗ Failed
rate_of_change_test	⊗ Failed
spike_test	S Failed

Sep 22, 3:38 PM

⊗ 3 ▲ 2 ⊘ 3

What's Next?

• Gulf of Maine Research Institute

- Move all legacy data pipelines into the new system
 - Retire messy old infrastructure
- Enhance pipeline resiliency
 - Accommodate data changes from providers
- Improve/standardize documentation
- Transition from ops to assets?
- Water level extravaganza!
- Drive additional services
 - CKAN
 - Model viewer



Why is our system called Sea Eagle?



• Gulf of Maine Research Institute

- System Inspiration from GLOS's Seagull
- I spent a month during initial deployment intentionally playing tour manager for a very lost Steller's Sea Eagle

