

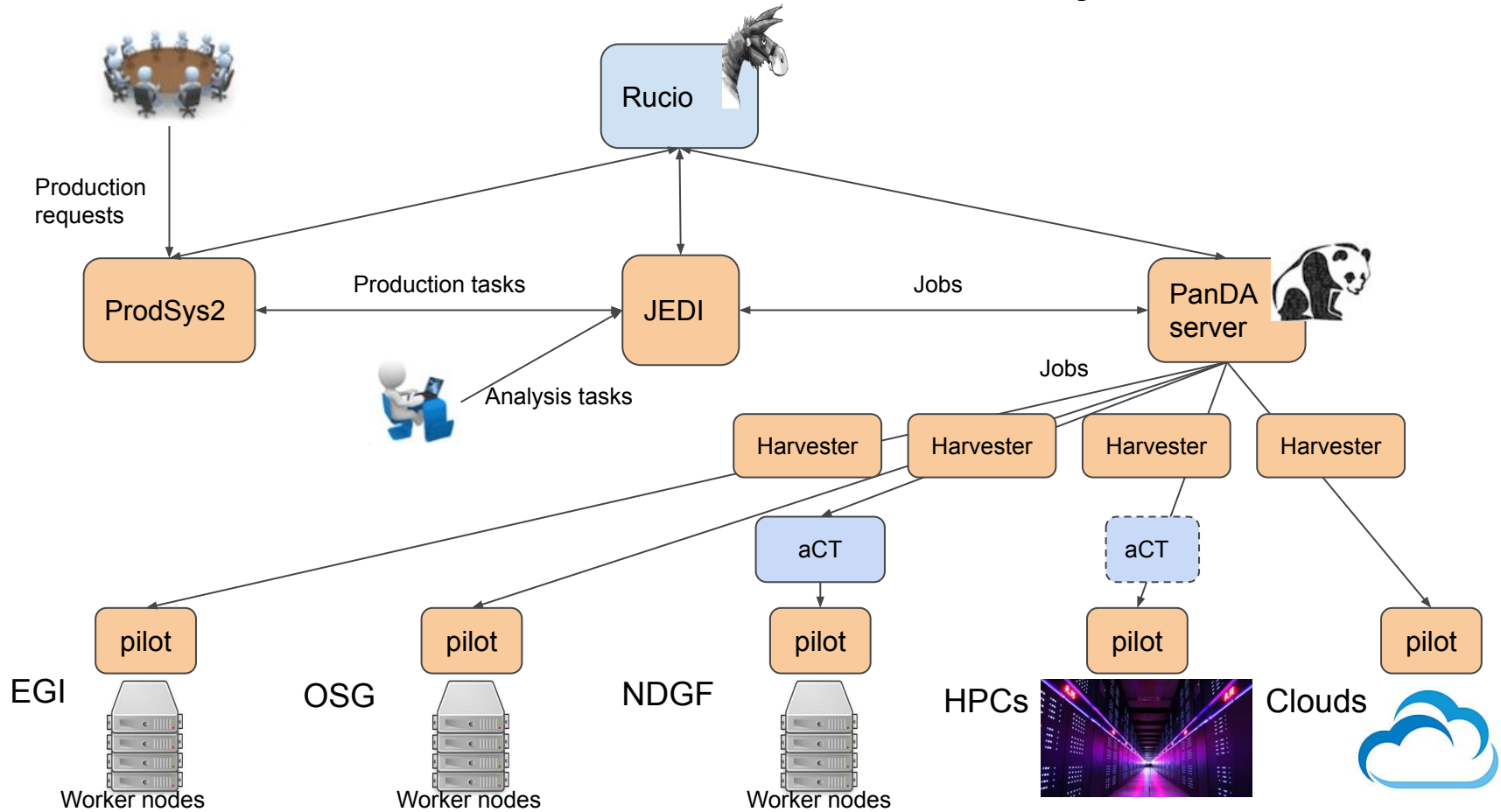
# Conscious Computing in ATLAS: Estimating Carbon Footprint for PanDA Workloads

Fernando Barreiro Megino

on behalf of the PanDA team

WLCG Env Sustainability Workshop, 12 December 2024

# PanDA: central orchestration of all ATLAS jobs



# Estimating job CO2 footprints through PanDA

- Overarching goal of estimating the CO2 footprint of the ATLAS experiment
  - See [Ben Bruers' ATLAS Week presentation](#)
- Feb 2023 Ben and Rod asked to add estimation for compute in PanDA

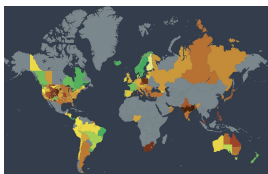
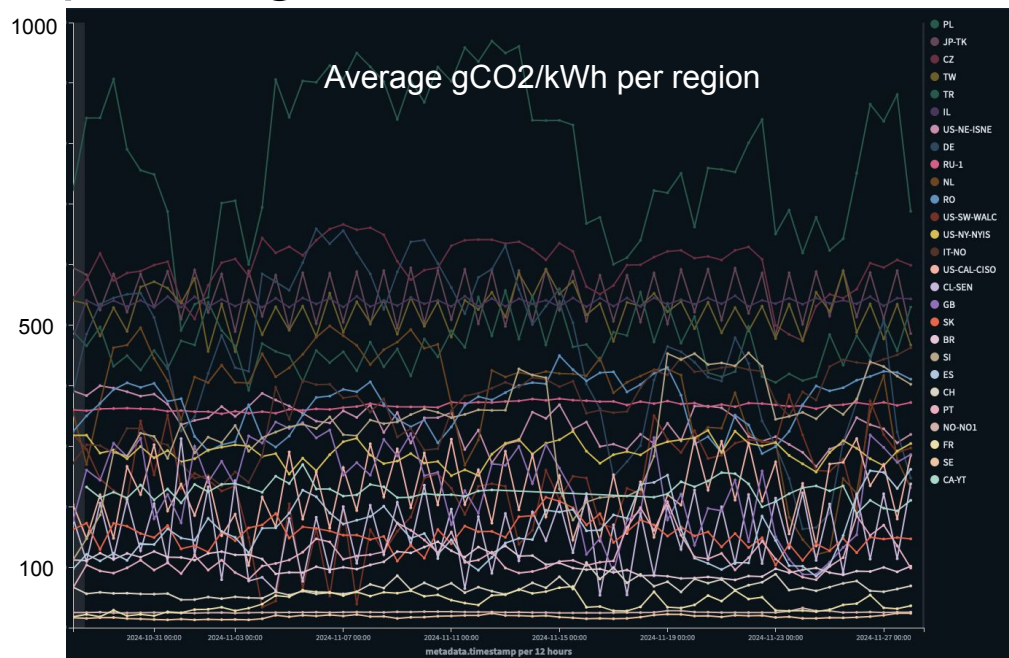
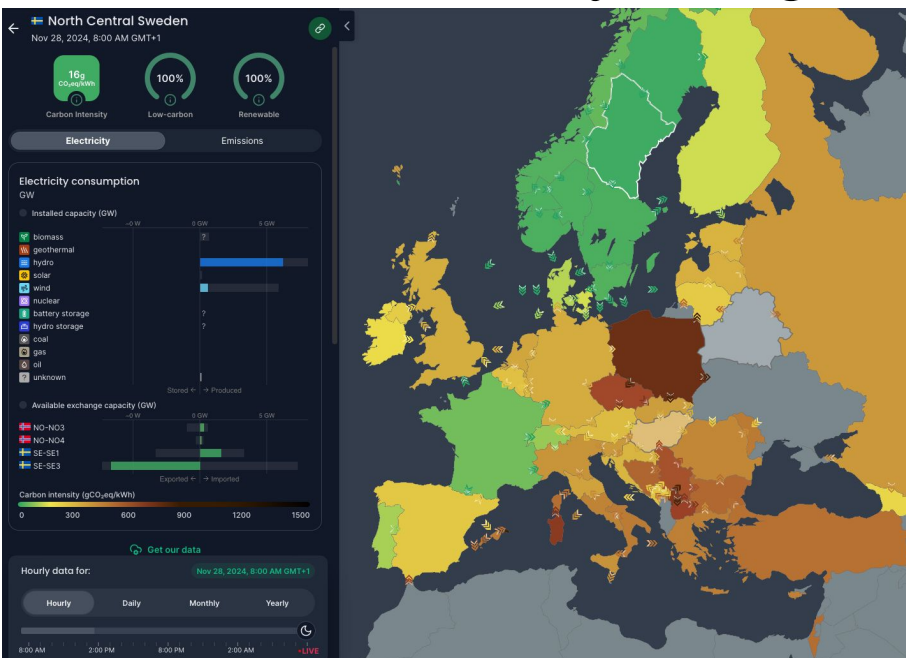


Image from  
<https://app.electricitymaps.com/map>



# Emission intensity of regional power grids



<https://app.electricitymaps.com/map/24h> (API: [CO2 signal](#))

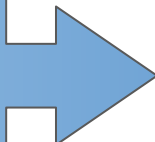
[CERN OpenSearch index](#)

- Data is available for many countries, particularly in Europe and the US
  - A few countries hosting ATLAS sites are missing
- In some cases data is published at regional level
- Emission intensities vary by region and fluctuate over time

# Emission intensities in the PanDA database

- Copy the **regional** emission intensity from the OpenSearch index
- Calculate **global** intensity as **weighted average** based on HS23 contributions

Regional intensities			gCO2/kWh	Weights (dynamically calculated)		
REGION	TIMESTAMP	VALUE		REGION	ROUND(SUM(JSS.HS)/TMP_TOTAL.TOTAL_HS,3)	
BR	29-NOV-24 09:30:03	95		BR	0.004	
CA-YT	29-NOV-24 09:30:04	216		CA-YT	0.047	
CH	29-NOV-24 09:30:05	61		CH	0.103	
CL-SEN	29-NOV-24 09:30:07	195		CN	0.002	
CZ	29-NOV-24 09:30:09	608		CZ	0.015	
DE	29-NOV-24 09:30:10	531		DE	0.127	
ES	29-NOV-24 09:30:11	181		ES	0.051	
FR	29-NOV-24 09:30:12	49		FR	0.067	
GB	29-NOV-24 09:30:14	187		GB	0.093	
IL	29-NOV-24 09:30:15	480		IL	0.008	
IT-NO	29-NOV-24 09:30:16	457		IT-NO	0.032	
JP-TK	29-NOV-24 09:30:17	589		JP-TK	0.027	
NL	29-NOV-24 09:30:18	449		NL	0.017	
NO-N01	29-NOV-24 09:30:19	25		NO-N01	0.005	
PL	29-NOV-24 09:30:21	796		PL	0.006	
PT	29-NOV-24 09:30:22	115		PT	0.002	
RO	29-NOV-24 09:30:23	352		RO	0.006	
RU-1	29-NOV-24 09:30:24	369		RU-1	0.004	
SE	29-NOV-24 09:30:26	23		SE	0.022	
SI	29-NOV-24 09:30:27	442		SI	0.17	
SK	29-NOV-24 09:30:28	154		SK	0.003	
TR	29-NOV-24 09:30:29	456		TR	0.003	
TW	29-NOV-24 09:30:30	497		TW	0.001	
US-CAL-CISO	29-NOV-24 09:30:31	338		US-CAL-CISO	0.003	
US-NE-ISNE	29-NOV-24 09:30:33	293		US-NE-ISNE	0.012	
US-NY-NYIS	29-NOV-24 09:30:34	271		US-NY-NYIS	0.045	
US-SW-WALC	29-NOV-24 09:30:35	356		US-SW-WALC	0.125	



Global intensities			gCO2/kWh
REGION	TIMESTAMP	VALUE	
GRID	29-NOV-24 09:30:35	315	

# Estimate gCO2 per job

$$n\_cores \times core\_power\_consumption \times \int_{starttime}^{endtime} emission\_intensity(t) dt$$

How power hungry the CC/HW is  
(default of 10W)

Carbon intensity of the power source  
over the duration of the job

gCO2/kWh

CH	16-FEB-23 10:30:33	93
CH	16-FEB-23 11:31:11	96
CH	16-FEB-23 12:31:01	86

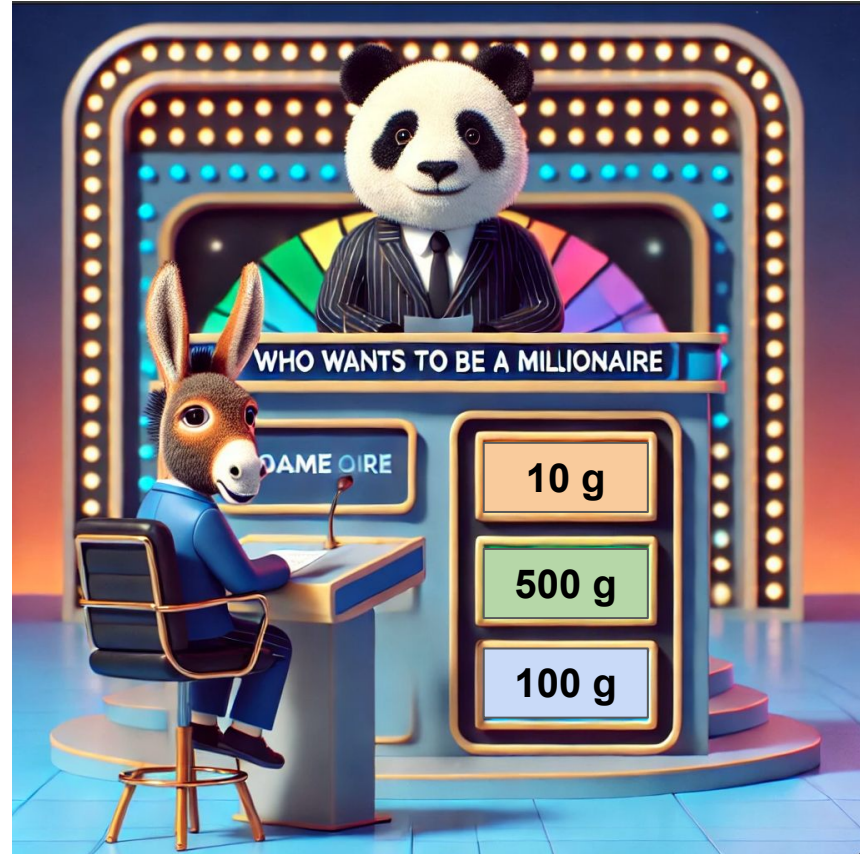
E.g. a 8 core job running at CERN from 11:00 to 12:30 would emit roughly

$$8 * 0.5h * 10W * 93 \text{ gCO2/kWh} + 8 * 1h * 10W * 96 \text{ gCO2/kWh} = 11.4 \text{ gCO2}$$

- The model is a first approximation for awareness
  - We are not pretending to provide precise results
- Sites can adjust their region and core power consumption in CRIC

# Job emission examples

What are the approximate CO<sub>2</sub> emissions for a 8-core, 24h job considering average emission intensity?



# Examples for heavy and light jobs

The CO2 estimations are stored in the database for each job entry and displayed in BigPanDA monitoring job pages

## Heavy jobs: multi-core, long duration

CORECOUNT	DURATION_HOURS	GCO2_REGIONAL	GCO2_GLOBAL
8	55.02	2382.98	1284.68
8	36.59	1073.84	851.93
8	35.8	1044.47	831.03
8	32.45	1042.72	755.94
8	35.52	1036.26	823.99
8	35.42	1034.78	822.07
8	59.3	973.01	1399.04
8	46.4	953.24	1073.75
8	45.53	939.75	1051.45
56	1.89	651.02	325.19
56	1.91	641.36	329.94
56	1.9	637.99	329.42
56	1.9	636.72	328.73
56	1.9	636.46	328.58
56	1.9	635.85	330.16

## Light jobs: single-core, short duration

CORECOUNT	DURATION_SECONDS	GCO2_REGIONAL	GCO2_GLOBAL
1	52	0.01	0.05
1	89	0.01	0.08
1	83	0.01	0.07
1	45	0.01	0.04
1	89	0.01	0.08
1	80	0.01	0.07
1	76	0.01	0.07
1	26	0.01	0.02
1	39	0.01	0.03
1	80	0.01	0.07
1	94	0.01	0.08
1	44	0.01	0.04
1	91	0.01	0.08
1	84	0.01	0.07
1	29	0.01	0.03



# Presentation of results and raising awareness


- Importance of the way the data is presented
- Users are presented with aggregates of **global** emissions (not regional)
  - Users typically don't select the region
  - Assigning jobs to a low-intensity region merely shifts other workloads to high-intensity regions
- The regional and global CO2 values are in the PanDA jobs table
  - This table is replicated to a few ElasticSearch/OpenSearch indexes
  - It is possible to make aggregated dashboards with CO2 emissions by activity, user, site
- Data is presented as approved by the ATLAS SW&C administration
  - I will only show examples that have been formally approved

# Increase awareness: task aggregation

## Task completion notification

JEDI notification for TaskID:42116242 (3/3 All Succeeded)

AA atlas-adc-panda-no-reply (no-reply address for PanDA) <atlas-adc-panda-no-reply@cern.ch> Today at 19:18  
To: Fernando Harald Barreiro Megjino



**Summary of TaskID: 42116242**

Detail	Value
Created	2024-11-18 18:12:14
Ended	2024-11-18 18:18:14.673831
Final Status	done

**Total Number of Inputs**

Category	Count
Succeeded	3
Failed	0
Cancelled	0
Total	3

### Estimated Carbon Footprint for the Task

Category	gCO2
Succeeded	0.14 gCO2
Failed	0 gCO2
Cancelled	0 gCO2
Total	0.14 gCO2

More details on estimation: [https://panda-wms.readthedocs.io/en/latest/advanced/carbon\\_footprint.html](https://panda-wms.readthedocs.io/en/latest/advanced/carbon_footprint.html)

### Panda Problem Reporting

- The eGroup for help request: [hn-atlas-dist-analysis-help@cern.ch](mailto:hn-atlas-dist-analysis-help@cern.ch)
- The Discourse forum for distributed computing help: <https://atlas-talk.web.cern.ch/c/distributed-computing-help>
- The JIRA portal for software bug: <https://its.cern.ch/jira/browse/ATLASPANDA>

## Task monitoring

42113766 task

Task ID	Type	Processing type	User	Nucleus	Status	N input files finished failed	N input events used	HS23s Expected Total done failed	Time stamps: created last modified	Cores	Priority: original current	Attempt	CO <sub>2</sub> total done failed
42113766	analy	panda-client-1.5.79-jedi-run	Arbitrary user		finished	110,322 (99%) 109,792 (0%) 530	14,339,794,846 (99.6%) 14,281,448,629	3.4G 2.2G 1.6G 553.6M	2024-11-18 12:28:00 2024-11-26 18:07:05	1	1001 1001	1	134kg 101kg 33kg

Task extra info Show jobs Task parameters and help Memory & walltime usage Other plots

States of jobs in this task [drop mode]

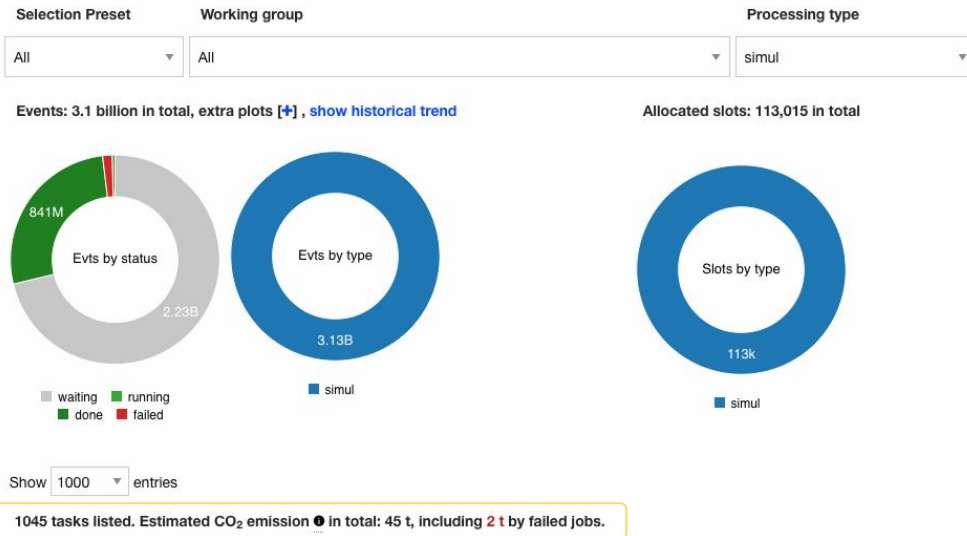
	pending	defined	waiting	assigned	throttled	activated	sent	starting	running	holding	transferring	merging	finished	failed	cancelled	closed
Build													76	9		1
Run													12808	460		
Merge													957	29		

# Increase awareness: user and activity aggregation

## “My BigPanDA”



## Running production tasks



# Conclusions

- Job footprint estimations available since Feb 2023
- They represent a rough, central estimation
- Handles available for sites to indicate more accurate core power consumption and region in CRIC
- The intent is to raise user awareness of computing emissions and encourage positive behaviour
  - Data added on request and as approved to monitoring
- PanDA does not carry out decisions based on CO2 emissions (e.g. broker to green regions)
  - Such decisions belong to the management

# Credits

- Idea and monit collector: Ben Bruers and Rod Walker
- Implementation in PanDA: Fernando Barreiro Megino
- Monitoring and accounting: Aleksandr Alekseev, Misha Borodin, Tatiana Korchuganova