A data-driven approach to carbon footprint reporting for High Performance Computing (HPC) at the University of Cambridge

Kell Johnston

Aims of the Project

- Provide an accurate evaluation of the energy used / carbon footprint of the HPC service
- Giving service users access to a dashboard showing their energy use / carbon footprint

How did we define Carbon Footprint

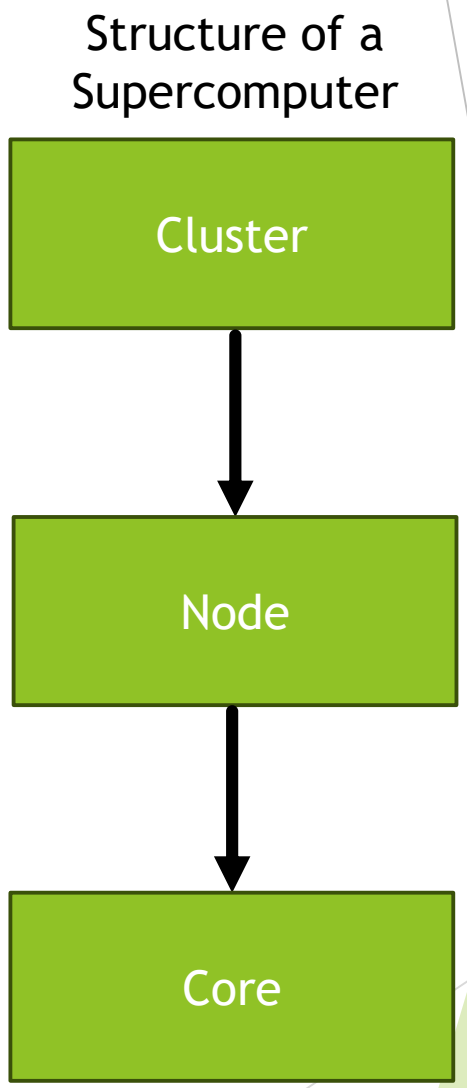
- Here Carbon footprint comes from 2 factors:
  - Emissions through use, where the use of energy from the power grid creates some Carbon dioxide
  - Embodied Carbon, the emissions created by manufacture, disposal and transportation of a product
- All “good” IT vendors can relay information about the Embodied Carbon associated with their products

Why is the information useful?

- HPC energy use is always increasing
- Currently HPC lacks carbon emissions feedback
- Lots of users are part of Net Zero targets.
- The data will feedback to users how much an increase in efficiency will reduce carbon emissions

HPC at UoC

- CSD3 (Cambridge Service for Data Driven Discovery) is the Tier 2 HPC Service offered by the University. It is comprised of 5 different clusters.
- Each Cluster can have several hundred nodes
- Nodes are “individual computers” each uniquely labelled and containing up-to 128 cores.
- Cores are a single processing unit, the smallest possible unit that a user could ask for.
- CSD3 utilises scheduling software called SLURM, which allows users to run jobs (individual programs)
- One Job != One Node
- One of clusters called “Ice-Lake” has 544 nodes, each node containing 76 cores meaning a possible 41344 simultaneous jobs on one cluster alone



What data do we have?

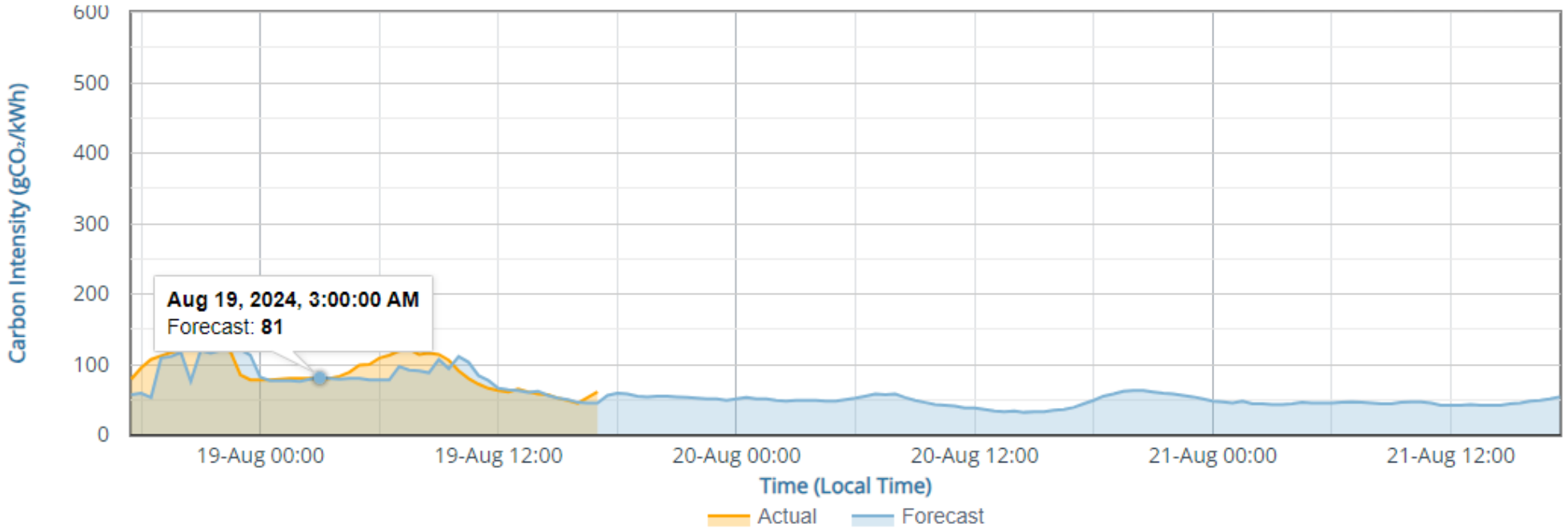
- SLURM Accounting Data
- Time Series Power Data
- Time Series Carbon Intensity Data
- Data on embodied carbon from Dell

Power Data


Power on gpu-q-46
2,400
2,200
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
0
Sat 15
06 AM
12 PM
06 PM
Jun 16
06 AM
12 PM
06 PM
Mon 17
06 AM
12 PM
06 PM
Tue 18
06 AM
12 PM
06 PM
Wed 19

Carbon Intensity

https://carbonintensity.org.uk/



SLURM Data

JobIDRaw	JobName	Partition	ElapsedRaw	Account	State	CPUTimeF	NodeList	User	AllocCPUS	AllocNodes	QOS	Start	End	Timelimit	Suspended
53822802	073d7cd5cb9de	cclake	878	33925368b054e0	COMPLETEI	7024	cpu-p-298	7dba682d5d5e	8	1	cpu1	2024-06-09T14:48:37	2024-06-09T15:03:15	06:00:00	00:00:00
53822803	073d7cd5cb9de	cclake	860	33925368b054e0	COMPLETEI	6880	cpu-p-298	7dba682d5d5e	8	1	cpu1	2024-06-09T14:48:37	2024-06-09T15:02:57	06:00:00	00:00:00
53822804	073d7cd5cb9de	cclake	1039	33925368b054e0	COMPLETEI	8312	cpu-p-298	7dba682d5d5e	8	1	cpu1	2024-06-09T14:48:37	2024-06-09T15:05:56	06:00:00	00:00:00
53822805	073d7cd5cb9de	cclake	1474	33925368b054e0	COMPLETEI	10318	cpu-p-607	7dba682d5d5e	7	1	cpu1	2024-06-09T14:48:37	2024-06-09T15:13:11	06:00:00	00:00:00
53822806	073d7cd5cb9de	cclake	192	33925368b054e0	FAILED	1920	cpu-p-299	7dba682d5d5e	10	1	cpu1	2024-06-09T14:48:37	2024-06-09T14:51:49	01:00:00	00:00:00
53822807	d217b8010b7d8	icelake-himem	3608	20c14bdd3659d3	TIMEOUT	137104	cpu-q-282	d22344c4929d	38	1	intr	2024-06-09T14:49:07	2024-06-09T15:49:15	01:00:00	00:00:00
53822917	c1f1b2f30f780c4	icelake	7	b300767227788e	COMPLETEI	532	cpu-q-304	31492ba50268	76	1	cpu1	2024-06-09T14:53:47	2024-06-09T14:53:54	01:00:00	00:00:00
53823134	c1f1b2f30f780c4	icelake	8	b300767227788e	COMPLETEI	608	cpu-q-287	31492ba50268	76	1	cpu1	2024-06-09T15:04:05	2024-06-09T15:04:13	01:00:00	00:00:00
53823246	d2f923ce202543	icelake	268	2776cecd034c85	COMPLETEI	268	cpu-q-10	74a2ed0f15a2c	1	1	cpu1	2024-06-09T14:59:00	2024-06-09T15:03:28	00:30:00	00:00:00
53823247	e4293a6ad3ede	icelake	313	2776cecd034c85	COMPLETEI	313	cpu-q-40	74a2ed0f15a2c	1	1	cpu1	2024-06-09T14:59:00	2024-06-09T15:04:13	00:30:00	00:00:00
53823248	202443488125e0	icelake	441	2776cecd034c85	COMPLETEI	441	cpu-q-132	74a2ed0f15a2c	1	1	cpu1	2024-06-09T14:59:00	2024-06-09T15:06:21	00:30:00	00:00:00
53823267	d58e2170d125af	cclake	908	4c60129f1023f3a	COMPLETEI	101696	cpu-p-[79-80]	2662e4fa132cc	112	2	sqos1	2024-06-09T14:56:16	2024-06-09T15:11:24	00:30:00	00:00:00
53823300	c5d0fd6ffde4e23	icelake-himem	28812	bfa8c7a63ef10bf	CANCELLED	8758848	cpu-q-[304-305	7e7de727a4ca	304	4	cpu1	2024-06-09T14:59:19	2024-06-09T22:59:31	1-12:00:00	00:00:00
53823307	3079ecf586ab7e	icelake	42670	6355f1b4e11ab7	COMPLETEI	5461760	cpu-q-[119,160	a92ac484b219	128	22	cpu2	2024-06-09T15:04:42	2024-06-10T02:55:52	12:00:00	00:00:00
53823410	ccc47f58514f9d7	icelake	1714	55647086daad13	COMPLETEI	54848	cpu-q-302	fe472afb98554	32	1	cpu2	2024-06-09T15:34:37	2024-06-09T16:03:11	12:00:00	00:00:00
53823430	088949f121dadb	ampere	72000	6224620299796d	TIMEOUT	2304000	gpu-q-17	8b2e510db95a	32	1	gpu1	2024-06-09T15:02:22	2024-06-10T11:02:22	20:00:00	00:00:00
53823437	6bc85d37245b58	icelake	3615	3b0e2161a3d456	TIMEOUT	7230	cpu-q-40	230c22ff8ee598	2	1	cpu2	2024-06-09T15:09:51	2024-06-09T16:10:06	01:00:00	00:00:00
53823438	daaf86ef181140	icelake	31	b300767227788e	COMPLETEI	2356	cpu-q-453	31492ba50268	76	1	cpu1	2024-06-09T15:09:00	2024-06-09T15:09:31	02:00:00	00:00:00
53823551	9582bc9b1d4e40	icelake	40148	56668a9abad164	COMPLETEI	73229952	cpu-q-[5,7,9,19	d54e406873f7f3	1824	24	dirac-lowu-cpu1	2024-06-09T21:56:21	2024-06-10T09:05:29	1-12:00:00	00:00:00
53823552	d217b8010b7d8	ampere	1354	2b35d581d6cb5c	CANCELLED	43328	gpu-q-17	872c27376c1e5	32	1	intr	2024-06-09T15:05:24	2024-06-09T15:27:58	01:00:00	00:00:00
53823565	b3497430c2c120	ampere	3813	318dc6ed61d427	COMPLETEI	488064	gpu-q-80	933d0fd1ddc55	128	1	gpu2	2024-06-09T15:05:55	2024-06-09T16:09:28	12:00:00	00:00:00
53823578	8b80eee620bb80	icelake-himem	0	cfddc2c1021b19	CANCELLED	0	None assigned	f9d5ab30f5822	0	0	cpu2		2024-06-09T15:06:31	10:00:00	00:00:00
53823579	088949f121dadb	ampere	222	fd7922ea8a996b	CANCELLED	7104	gpu-q-21	192a48fa9f806	32	1	gpu1	2024-06-09T15:06:26	2024-06-09T15:10:08	1-11:00:00	00:00:00
53823580	0c72058cb79c9a	icelake-himem	597	84522b7b10ecf93	COMPLETEI	19104	cpu-q-349	bff6d77257239	32	1	cpu2	2024-06-09T15:07:42	2024-06-09T15:17:39	10:00:00	00:00:00
53823587	5c4cb1606e1d2e	sapphire	129616	70edb26e8ae31a	TIMEOUT	72584960	cpu-r-[9,18,21,	4757fe99223ca	560	5	t2-cpu1	2024-06-10T08:24:04	2024-06-11T20:24:20	1-12:00:00	00:00:00
53823620	c1f1b2f30f780c4	icelake	8	b300767227788e	COMPLETEI	608	cpu-q-287	31492ba50268	76	1	cpu1	2024-06-09T15:14:12	2024-06-09T15:14:20	01:00:00	00:00:00
53823736	6603bf4b986b55	ampere	92818	2a40cf9a1dd5b1	COMPLETEI	2970176	gpu-q-21	c6d7d845ecc0	32	1	gpu1	2024-06-09T15:11:30	2024-06-10T16:58:28	1-02:00:00	00:00:00
53823662	6603bf4b986b55	ampere	93015	2a40cf9a1dd5b1	COMPLETEI	2976480	gpu-q-33	c6d7d845ecc0	32	1	gpu1	2024-06-09T15:11:30	2024-06-10T17:01:45	1-02:00:00	00:00:00

A Simplified Example:

JobIDRaw	Partition	ElapsedRaw	State	NodeList	AllocCPUS	AllocNodes	Start	End
1	icelake	43200	TIMEOUT	cpu-q-56	76	1	2024-06-24T20:56	2024-06-25T08:56:00
2	icelake	43200	TIMEOUT	cpu-q-48	38	1	2024-06-24T22:02	2024-06-25T10:02:00
3	icelake	43260	TIMEOUT	cpu-q-48	38	1	2024-06-24T22:40	2024-06-25T10:41:00
4	icelake	20640	FAILED	cpu-q-[188,217,258]	114	3	2024-06-24T18:03	2024-06-24T23:47:00
5	icelake	18000	COMPLETED	cpu-q-[188,217,258]	114	3	2024-06-24T21:33	2024-06-25T02:33:00
6	login-epicov	36000	TIMEOUT	cpu-b-51	8	1	2024-06-25T02:45	2024-06-25T12:45:00
7	epid	36000	COMPLETED	cpu-p-157	32	1	2024-06-25T11:06	2024-06-25T21:06:00

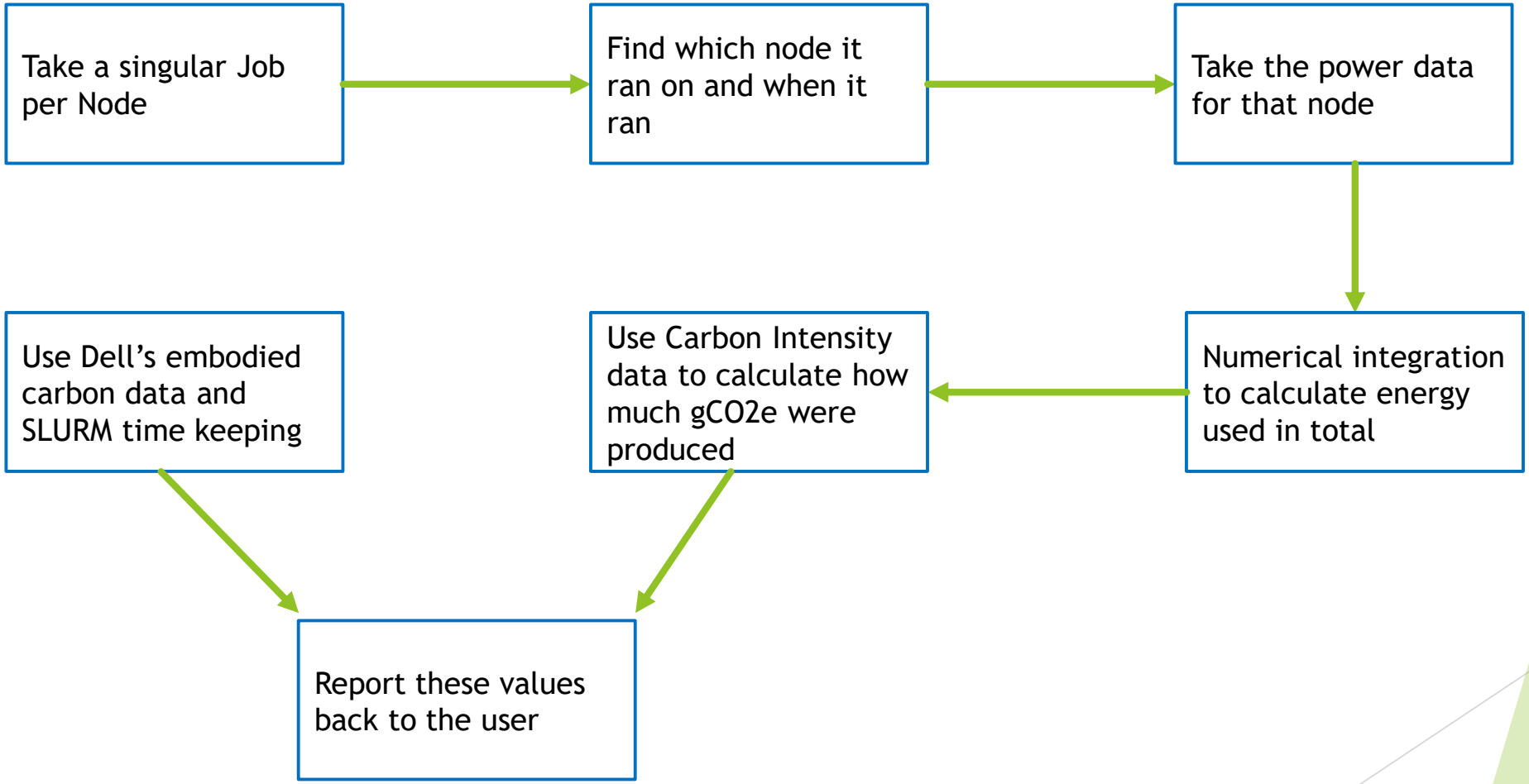
- Which node(s)
- When did it start and end
- Who submitted it
- How many Nodes and Cores were used?

Explode to Jobs per Node

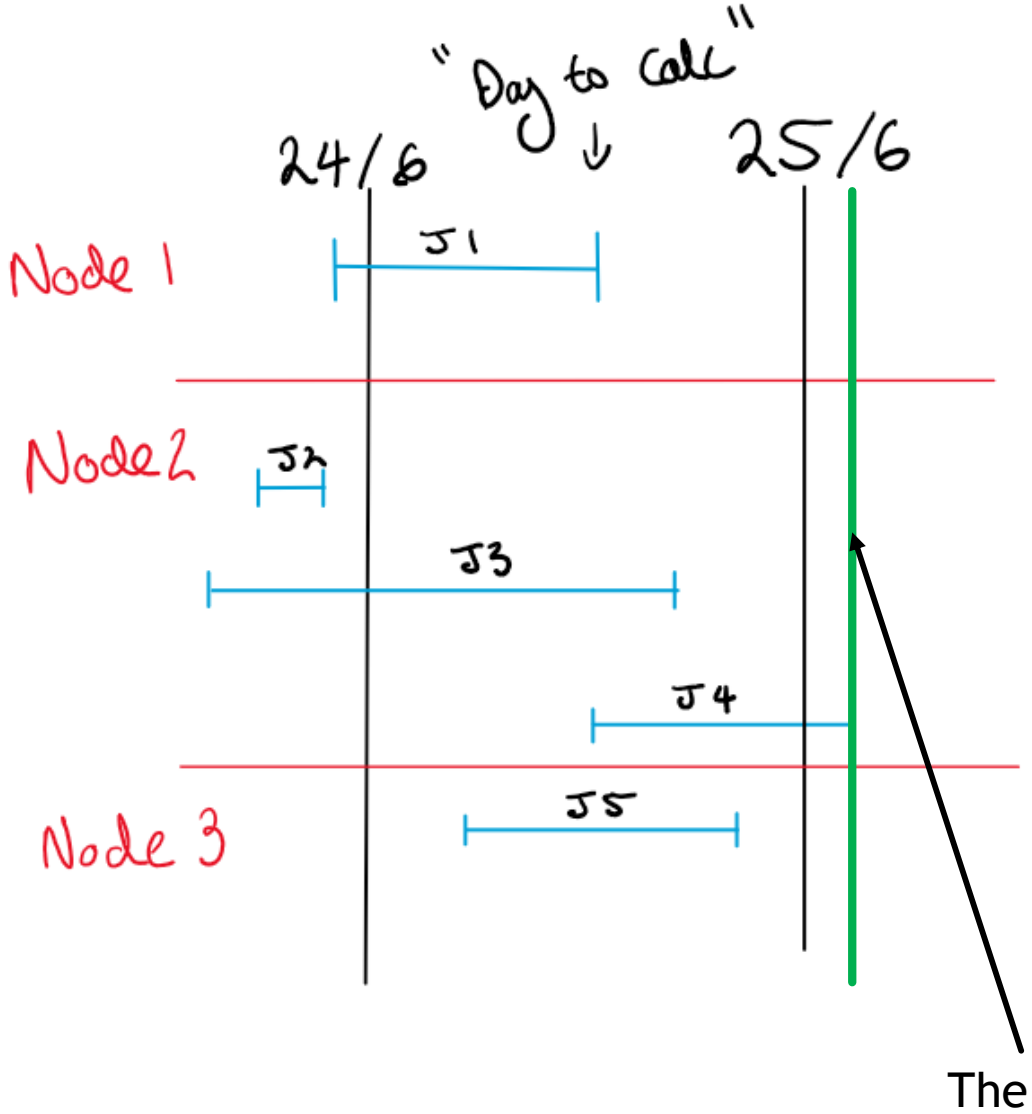
- Cannot handle multi-node jobs in their current state
- “Explode” the multi-node jobs into jobs per node

JobIDRaw	Partition	ElapsedRaw	State	NodeList	AllocCPUS	AllocNodes	Start	End
6	login-epicov	36000	TIMEOUT	cpu-b-51	8	1	25/06/2024 02:45	25/06/2024 12:45
7	epid	36000	COMPLETED	cpu-p-157	32	1	25/06/2024 11:06	25/06/2024 21:06
2	icelake	43200	TIMEOUT	cpu-q-48	38	1	24/06/2024 22:02	25/06/2024 10:02
3	icelake	43260	TIMEOUT	cpu-q-48	38	1	24/06/2024 22:40	25/06/2024 10:41
1	icelake	43200	TIMEOUT	cpu-q-56	76	1	24/06/2024 20:56	25/06/2024 08:56
4	icelake	20640	FAILED	cpu-q-188	114	3	24/06/2024 18:03	24/06/2024 23:47
4	icelake	20640	FAILED	cpu-q-217	114	3	24/06/2024 18:03	24/06/2024 23:47
4	icelake	20640	FAILED	cpu-q-258	114	3	24/06/2024 18:03	24/06/2024 23:47
5	icelake	18000	COMPLETED	cpu-q-188	114	3	24/06/2024 21:33	25/06/2024 02:33
5	icelake	18000	COMPLETED	cpu-q-217	114	3	24/06/2024 21:33	25/06/2024 02:33
5	icelake	18000	COMPLETED	cpu-q-258	114	3	24/06/2024 21:33	25/06/2024 02:33

Outlining the Process



Problems with that approach:



This is not meant to match the previous example and is a separate illustration

The current time

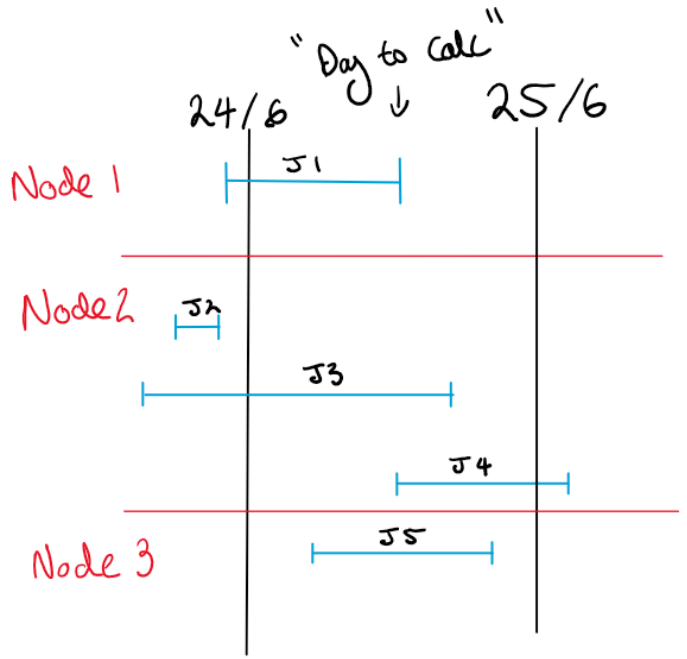
Exclusive, Overlapping and Lonely

Job 1: Uses 100% of the possible cores on the node(s)

Job 3: Uses 50% of the possible cores on the node and overlaps with Jobs 2 and 4

Job 5: Uses 80% of the possible cores on the node(s) but doesn't overlap with any Jobs

#Cores allocated = #Total Number of cores in used Nodes



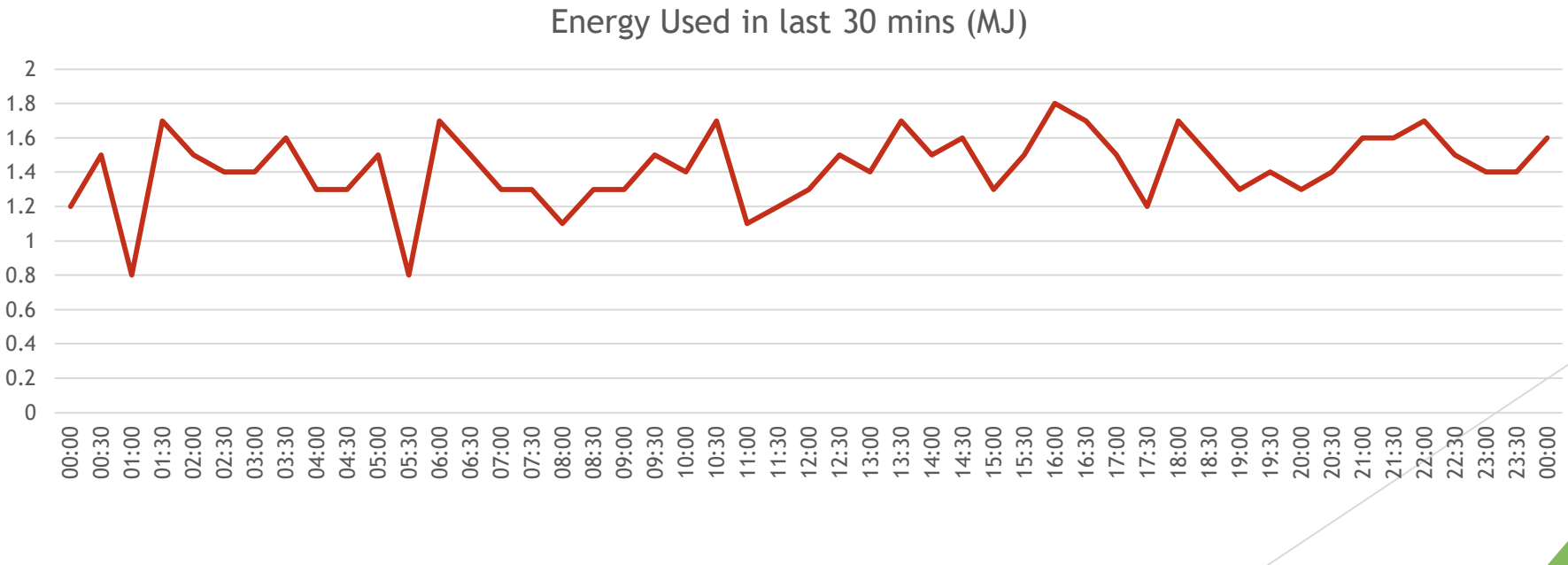
Return to Example

JobIDRaw	Partition	ElapsedRaw	State	NodeList	AllocCPUS	AllocNodes	Start	End
1	icelake	43200	TIMEOUT	cpu-q-56	76	1	2024-06-24T20:56	2024-06-25T08:56:00
2	icelake	43200	TIMEOUT	cpu-q-48	38	1	2024-06-24T22:02	2024-06-25T10:02:00
3	icelake	43260	TIMEOUT	cpu-q-48	38	1	2024-06-24T22:40	2024-06-25T10:41:00
4	icelake	20640	FAILED	cpu-q-[188,217,258]	114	3	2024-06-24T18:03	2024-06-24T23:47:00
5	icelake	18000	COMPLETED	cpu-q-[188,217,258]	114	3	2024-06-24T21:33	2024-06-25T02:33:00
6	login-epicov	36000	TIMEOUT	cpu-b-51	8	1	2024-06-25T02:45	2024-06-25T12:45:00
7	epid	36000	COMPLETED	cpu-p-157	32	1	2024-06-25T11:06	2024-06-25T21:06:00

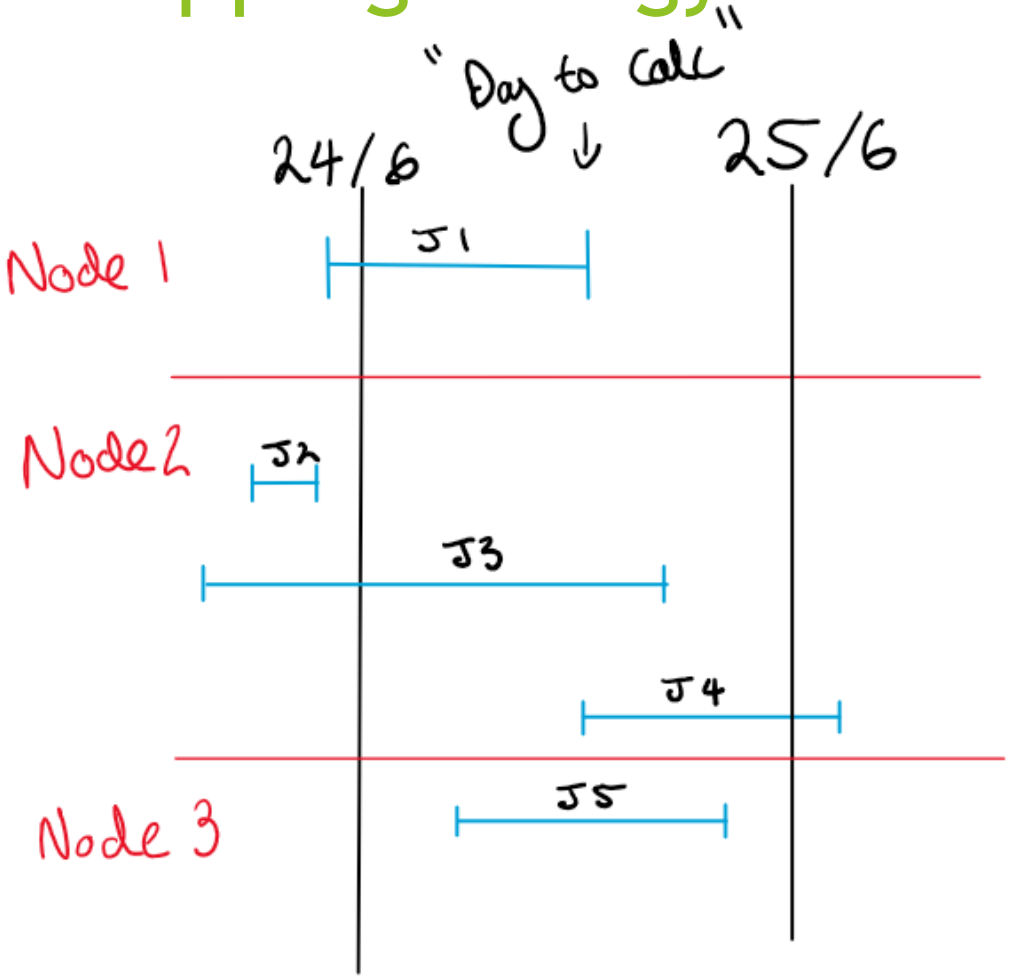
On an Icelake Node there are 76 possible cores.

On an EPID Node there are 64 possible cores.

Exclusive and Lonely Energy



Overlapping Energy

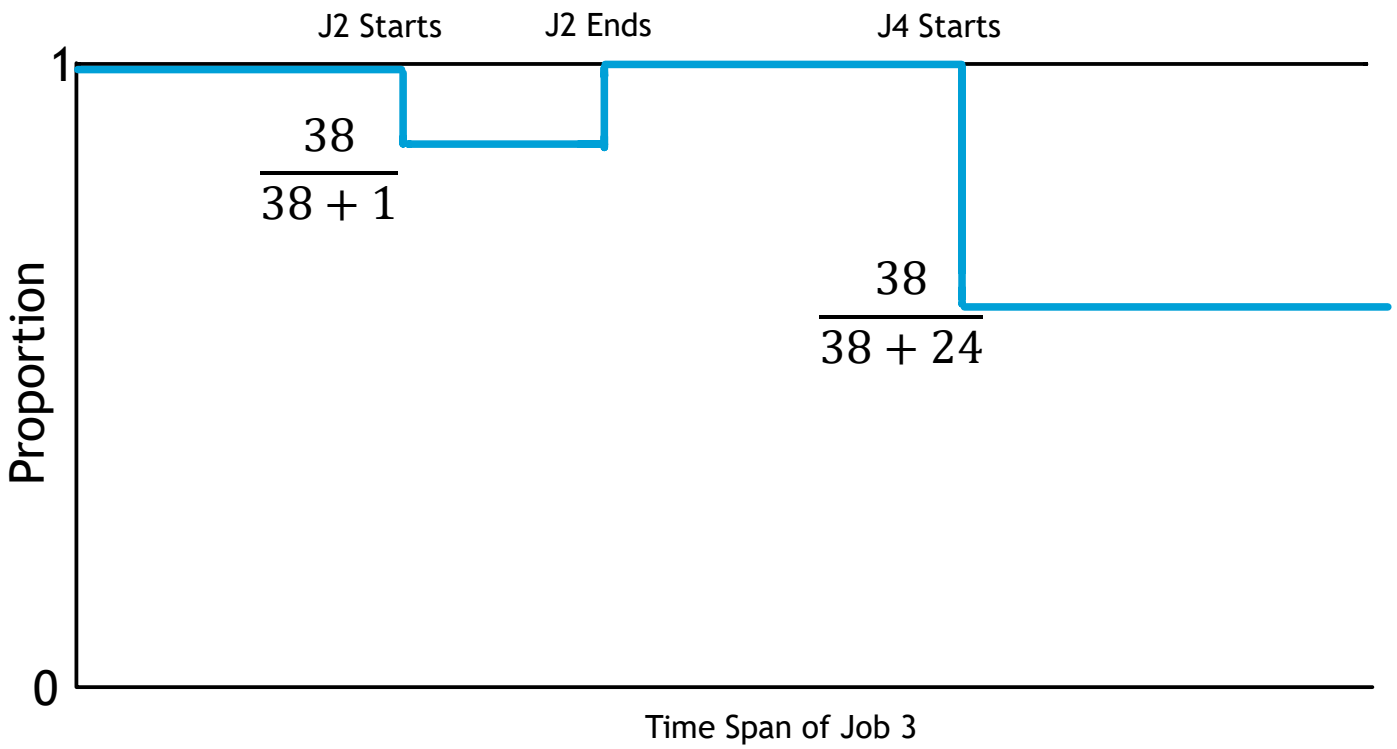


Job 2: Uses 1 core of a possible 76

Job 3: Uses 38 cores of a possible 76

Job 3: Uses 24 cores of a possible 76

Overlapping Energy



Results

JobIDRaw	Partition	ElapsedRaw	NodeList	Start	End	JobNodeEnerg	JobNodeCar	EmbCarbon
1	icelake	43200	['cpu-q-56']	24/06/2024 20:56	25/06/2024 08:56	25296390	955.203274	442.6902433
2	icelake	43200	['cpu-q-48']	24/06/2024 22:02	25/06/2024 10:02	16872240	647.727514	221.3451216
3	icelake	43260	['cpu-q-48']	24/06/2024 22:40	25/06/2024 10:41	16916760	656.690874	221.6525454
5	icelake	18000	['cpu-q-188', 'cp	24/06/2024 21:33	25/06/2024 02:33	31097503.5	994.007082	276.681402
6	login-epicov	36000	['cpu-b-51']	25/06/2024 02:45	25/06/2024 12:45			
7	epid	36000	['cpu-p-157']	25/06/2024 11:06	25/06/2024 21:06	7358400	298.122704	

- Job 4 is missing?
- Job 6 has no data?
- Job 7 is missing embodied carbon data?

Embodied Carbon

- There are several creators of embodied carbon:
  - Manufacture
  - Transport
  - Disposal

The following equation gives amount of embodied carbon for that job

$$\frac{\text{Proportion of each Node} \times \text{Number of Nodes Used} \times \text{"}Length\ of\ program\ (s)\text{"}}{(Number\ of\ Seconds\ in\ 5\ years)} \times \text{Embodied carbon created per Node}$$

User Centred Dashboard

User Input Parameters

Select the appropriate date-range:

2024-06-15 2024-06-18

2024-06-15 2024-06-18

Which user are you?

7495df95a5d322ac1fe79…

Summary Data:

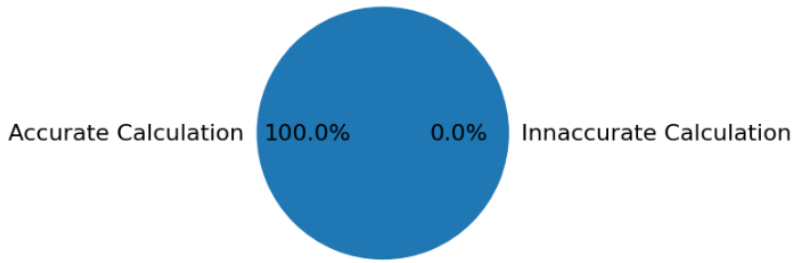
A summary of the data about the types of jobs that were submitted, including which partitions (i.e. cluster) the jobs ran on and whether the job successfully ran or not:

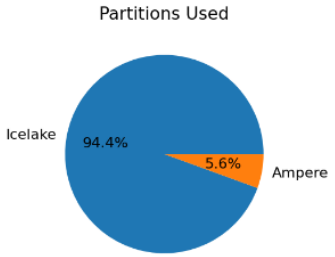
Total Energy (J) and gCO2 produced:

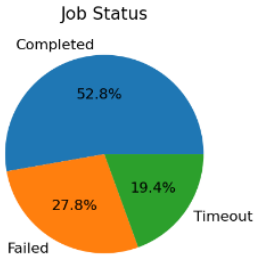
Number of Jobs	Total Energy	Total gCO2 through use	Total Embodied Carbon used gCO2	Total Carbon footprint gCO2
36	701,264,821.4766	18,583.8313	9,487.1037	28,070.935

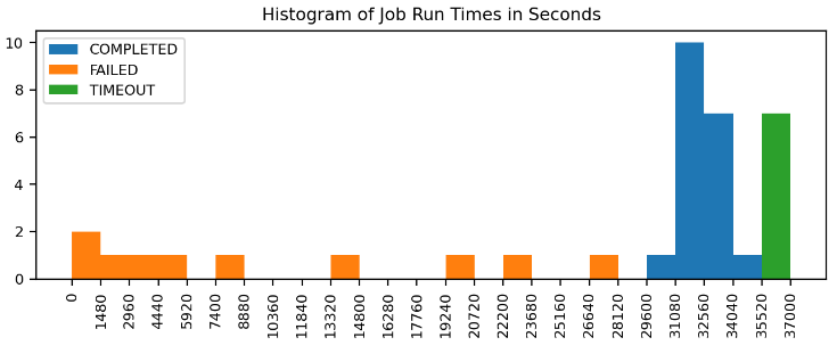
	ElapsedRaw	JobNodeEnergy	JobNodeCarbon	EmbCarbon	TotalCarbon
count	36	36	36	36	36
mean	27,145.8056	19,479,578.3744	516.2175	263.5307	779.7482
min	9	2,479.5	0.1211	0.0143	0.1353
50%	32,065	23,123,982.8825	563.6749	311.2908	885.7546
max	36,021	26,379,369.6956	870.8154	349.6962	1,220.4339

Proportion of Jobs with accurate calculations:









Energy Data:

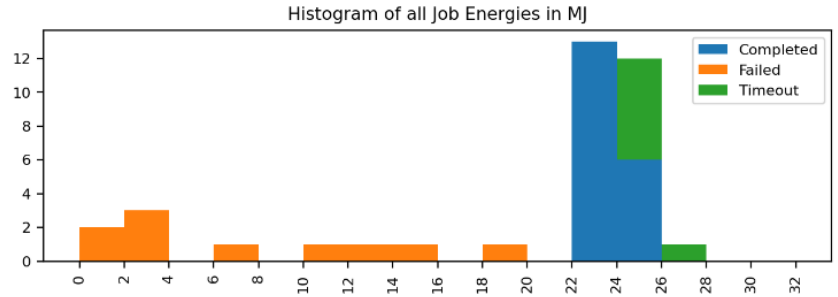
As a user the energy created by a job is a more accurate representation of efficiency and optimisation of your programs This is because there is unnacountable fluctuations in the amount of gCO2e/KwH produced by the power grid, so the same program ran at different times could give considerable different Carbon values, but would give a similar value for energy.

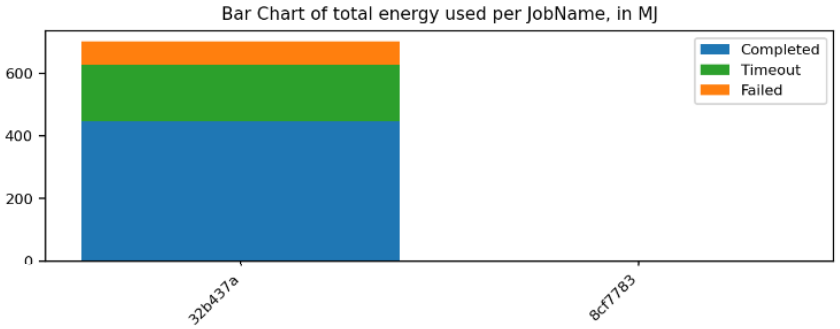
Job Name specific information:

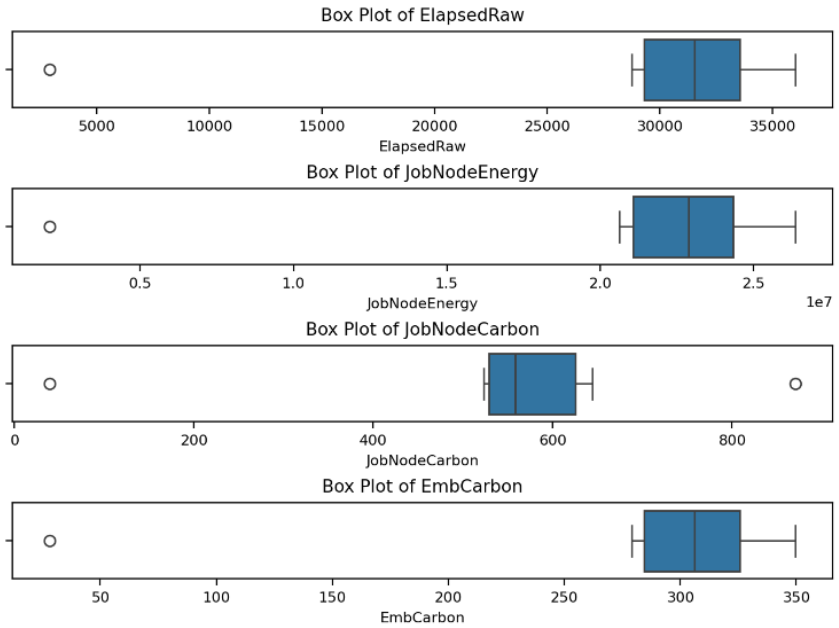
Which Job-Name do you want to see energy data for?

32b437a8759b2ac02cc0c67a3b5ebd6d6f482c3b9439698baf1507857da3e2ef

JobIDRaw	Partition	State	ElapsedRaw	JobNodeEnergy	JobNodeCarbon	EmbCarbon	TotalCarbon	Start
53,361,627	icelake	TIMEOUT	36,021	25,478,220.5796	570.5733	349.6962	920.2695	2024-06-14 19:5(
53,361,628	icelake	TIMEOUT	36,020	25,497,107.7215	570.9652	349.6865	920.6517	2024-06-14 19:51
53,361,629	icelake	TIMEOUT	36,020	25,556,361.5349	572.2772	349.6865	921.9637	2024-06-14 19:51
53,361,630	icelake	TIMEOUT	36,020	25,449,849.7622	570.0977	349.6865	919.7842	2024-06-14 19:51
53,361,631	icelake	FAILED	19,387	13,728,060.3158	312.0052	188.2113	500.2165	2024-06-14 19:51
53,361,634	icelake	COMPLETED	33,855	23,926,257.9405	534.9606	328.6684	863.629	2024-06-14 20:08
53,361,635	icelake	FAILED	4,319	3,043,066.5263	63.1968	41.9294	105.1261	2024-06-15 01:19
53,361,636	icelake	COMPLETED	32,068	22,680,850.1095	464.0722	311.32	775.3922	2024-06-15 02:39
53,361,640	icelake	COMPLETED	34,028	24,038,675.6293	557.2521	330.3479	887.6	2024-06-15 05:33
53,361,641	icelake	COMPLETED	32,062	22,626,053.269	523.0016	311.2617	834.2633	2024-06-15 05:51







Metadata for JobName: 32b437

	ElapsedRaw	JobNodeEnergy	JobNodeCarbon	EmbCarbon
count	34	34	34	34
mean	28,742.0882	20,625,266.3326	546.5752	279.0316
std	9,890.4461	7,136,795.2905	208.8489	96.0176
min	2,892	2,036,066.7895	38.9402	28.0759
25%	30,943.25	22,457,387.9792	523.3589	300.4008
50%	32,111.5	23,332,141.1817	570.3355	311.7423
75%	34,031.75	24,678,397.6782	643.9895	330.3843
max	36,021	26,379,369.6956	870.8154	349.6962

What next?

- Finishing the user focused dashboard
- Visualize metrics about the data across the whole of CSD3, for an internal report. Letting the UIS view data aggregated by cluster or department that uses the service.
- Lots of code commenting
- Considerations for production such as resourcing, automation and implementation

Any Questions?