Allocating Carbon Costs to Computing Payloads across Heterogeneous Infrastructures.

Measurement & Allocation

NetZero: The Big Picture

IRIS-CMP Allocation Models

Testing Models

Choosing between models

Final Thoughts

Speaker: Dr Alex Owen

r.a.owen@qmul.ac.uk

Queen Mary University of London

CHEP 2024-10

Measuring Data Centre Carbon Costs

IRISCAST

Audit of Carbon Costs

https://doi.org/10.5281/zenodo.7692451

https://doi.org/10.1051/epjconf/202429507015

Allocating Carbon Costs to Payloads

iris CMP

Alex Owen (QMUL)
Jon Hays (QMUL)

Daohai Li (QMUL)
Alex DIbbo (STFC)

https://doi.org/10.5281/zenodo.10966001

Queen Mary University of London

GridPP

UKRI Science and Technology Facilities Council | Scientific Computing

Allocating Carbon Costs to Payloads

iris CMP

https://doi.org/10.5281/zenodo.10966001

Scope 1 Emissions
Burning stuff

Eg Backup Generators

Scope 2 Emissions
Using the kit
Active Carbon

→ Operational Carbon

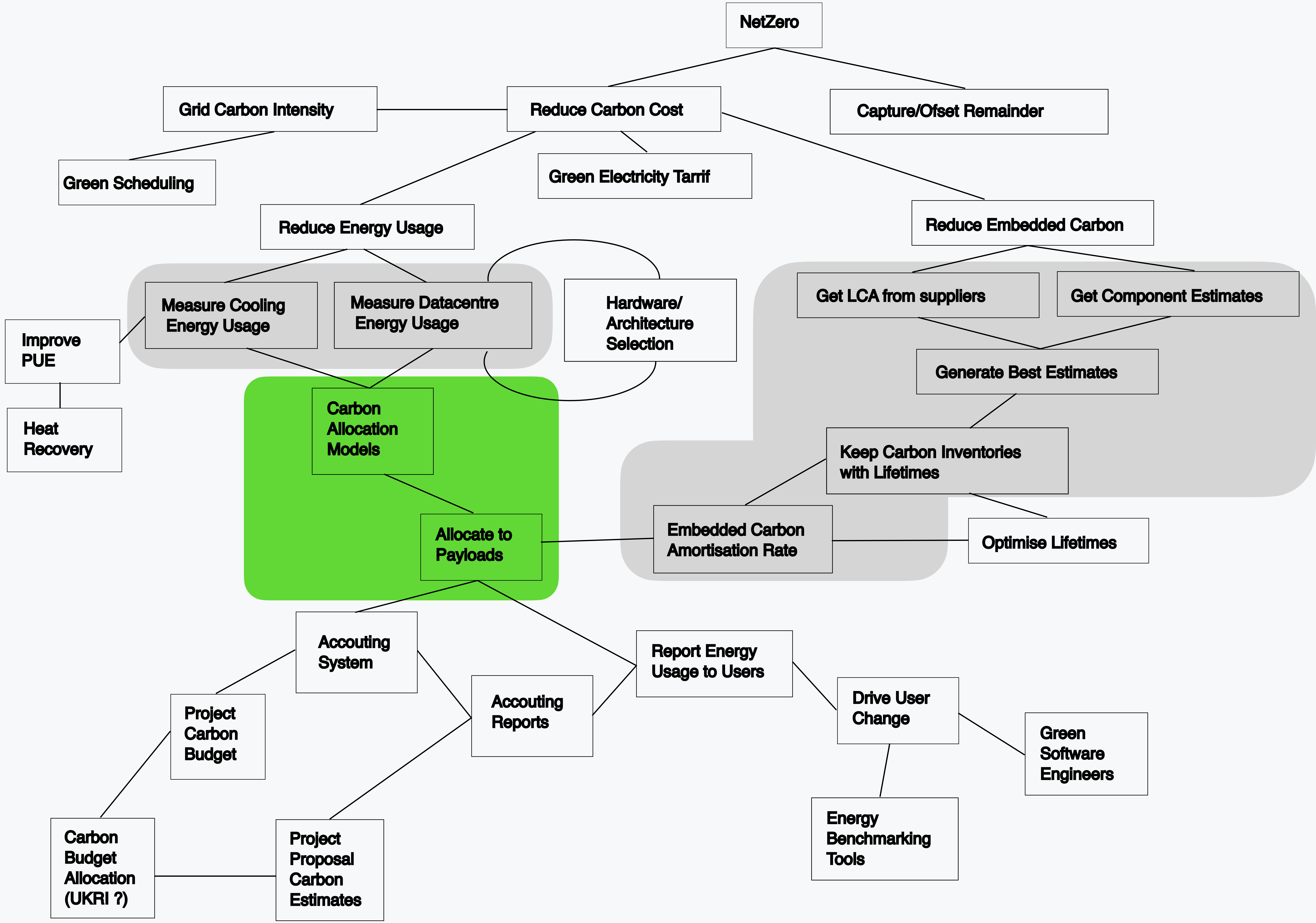
Like OpEX

Scope 3 Emissions
Obtaining the kit
Embedded Carbon

→ Capital Carbon

Like CapEX

NetZero The Big Picture



Based on figure in IRIS-CMP Report

iris Carbon Mapping Project

	Scope 2 – Energy	Scope 3 – Carbon
Payload	$E_p = E_f^t \cdot \frac{R_p}{R_f} \cdot \frac{t_p}{t}$	$C_{ep} = \frac{R_p}{R_f} \cdot t_p \cdot Q_{ef}$ Where: $Q_{ef} = \sum_{x=1}^{items} \frac{C_{ex}}{T_x}$
Idle	$E_{idle}^t = E_f^t - \sum_{p=1}^{payloads} Ep$	$C_{e\ idle}^t = t \cdot Q_{ef} - \left(\sum_{p=1}^{payloads} C_{ep} \right)$

Table 1: Summary of the Simple Payload Model showing allocations of Scope 2 energy and Scope 3 carbon to user payloads and the remaining idle allocation to the provider.

Apportion by Real Time

Simple Payload Model

Input	Description
E_f^t	Facility Energy usage over an accounting period (including cooling) could be estimated from PDU readings multiplied by PUE
t	Duration of accounting period
t_p	Elapsed time of a payload (Wall clock)
R_p	Resource slots allocated to job (eg CPU's)
R_f	Total slots available at facility
C_{ex}	Inventory Entry: Embedded carbon of each item x in facility
T_x	Inventory Entry: expected lifetime of each item x in facility

Table 2: Summary of the inputs needed to evaluate the Simple Payload Model.

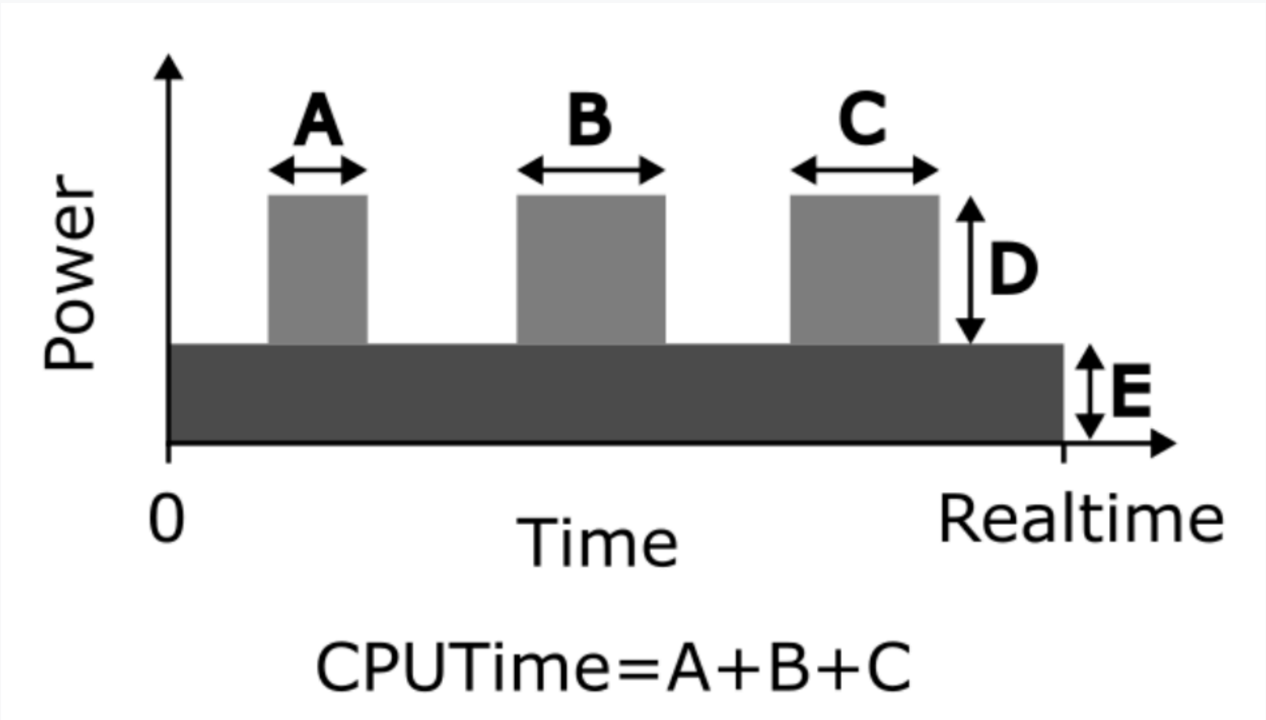
iris Carbon Mapping Project

	Scope 2 - Energy	Scope 3 - Carbon
Payload	$E_p = P_f^{idle} \cdot \frac{R_p}{R_f} \cdot t_p + P_{slot}^{CPU} \cdot t_p^{CPU}$ Where: $P_{slot}^{CPU} = \frac{E_f^t - P_f^{idle} \cdot t}{t_f^{CPU}}$	$C_{ep} = \frac{R_p}{R_f} \cdot t_p \cdot Q_{ef}$ Where: $Q_{ef} = \sum_{x=1}^{items} \frac{C_{ex}}{T_x}$
Idle	$E_{idle}^t = E_f^t - \sum_{p=1}^{payloads} Ep$	$C_{e\ idle}^t = t \cdot Q_{ef} - \left(\sum_{p=1}^{payloads} C_{ep} \right)$

Table 3: Summary of the Enhanced Payload Model showing allocations of Scope 2 energy and Scope 3 carbon to user payloads and the remaining idle allocation to the provider

Know your idle power?

Know your CPUtime?



Input	Description
E_f^t	Facility Energy usage over an accounting period (including cooling) could be estimated from PDU readings multiplied by PUE
P_f^{idle}	Idle power draw of the facility (including cooling) could be estimated from PDU readings during an idle period multiplied by PUE
t	Duration of accounting period
t_f^{CPU}	Total CPUtime delivered by the facility during the accounting period.
t_p	Elapsed time of a payload (Wall clock)
t_p^{CPU}	CPUtime of a payload
R_p	Resource slots allocated to job (eg CPU's)
R_f	Total slots available at facility
C_{ex}	Inventory Entry: Embedded carbon of each item x in facility
T_x	Inventory Entry: expected lifetime of each item x in facility

Table 4: Summary of the inputs needed to evaluate the Enhanced Payload Model.

Enhanced Payload Model

iris Carbon Mapping Project

Testing the Payload Models for Batch

Idle Power →

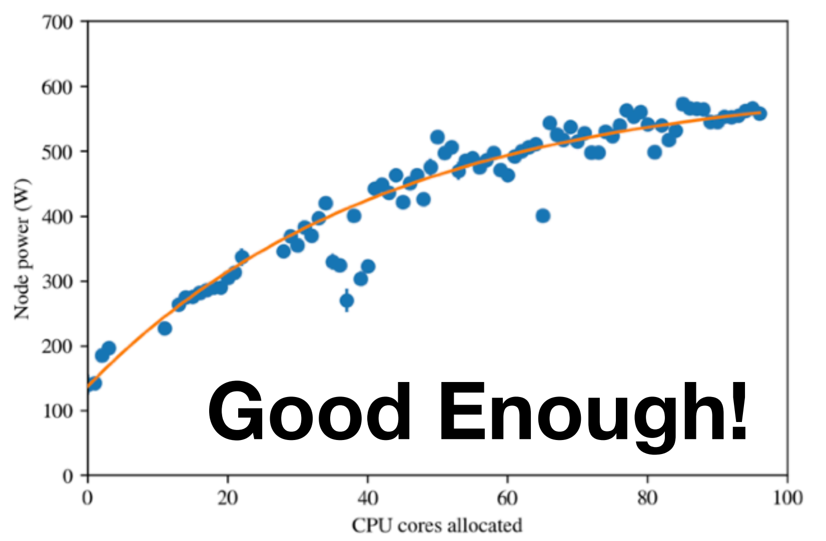


Figure 2: Node power P_N as a function of allocated CPU cores in a node N_{CA}. The Blue dots and error bars show the mean power and std-deviation for each discrete N_{CA}. Regression line y-intercept gives the idle power of a node $P_{idle} \approx 137.1W$.

User	Simple Payload Model kWh	Enhanced Payload Model kWh
prdatl	1204.79	1191.95
pillhcb	159.08	242.24
pilcms	76.83	71.28
pilatl	48.86	51.58
Pilmoe	10.75	16.86
Pildune	2.46	0.61
Others	0.08	0.04
Sub total	1502.86	1574.57
Idle(provider)	94.14	22.43
Total	1597	1597

Table 10: Results of evaluating the Simple and Enhanced Payload models on QMUL batch payloads the 24 hour period of 2024-03-07.

Input	Value	Slurm name	Description
E_f^t	1597 kWh	-	Facility Energy usage. In this four rack example the PDU cumulative energy readings were used to calculate this.
P_f^{idle}	16.45 kW	-	Idle power draw of the facility. In this example the 137.1W per node was multiplied by 120 nodes.
t	86400 s	-	Duration of accounting period. In this case 24 hours.
t_f^{CPU}	-	$\sum TotalCPU$	Total CPUtime delivered by the facility during the accounting period. Sum of the TotalCPU figures for all payloads
t_p	-	Elapsed	Elapsed time of a payload (Wall clock)
t_p^{CPU}	-	TotalCPU	CPUtime of a payload
$Slots_p$	-	AllocCPUS	Resource slots allocated to job (eg CPU's)
$Slots_f$	11520	-	Total slots available at facility. In this case 120 nodes with 96 cores each.

Table 9: Measured and derived constants and Slurm accounting data names used to evaluate the payload models for QMUL batch payloads.

Works for Batch!

iris Carbon Mapping Project

Testing the Payload Models for Cloud

Works for Cloud too!

User	Simple Payload Model kWh
Project 1	51.51
Project 2	31.52
Project 3	25.07
Project 4	18.22
Project 5	17.61
Project 6	12.89
Others	94.00
Sub total	250.82
Idle(provider)	173.44
Total	424.26

Input	Value	Prometheus name	Description
E_f^t	424.26 kWh	-	Facility Energy usage, derived from "node_hwmon_power_average_watt" and our accounting period t on all nodes.
t	72000 seconds	-	Duration of accounting period. In this case 20 hours.
t_p	-	-	Elapsed time of a VM (Wall clock) during our accounting period, as inferred by the VM's "launched_at" and "terminated_at" time from OpenStack.
R_p	-	openstack_nova_vcpus_used	Resource slots allocated to VM (eg CPU's)
R_f	?	openstack_nova_vcpus_available	Total slots available at facility. In this case number of all vcpus on all the nodes.

Table 13: Measured and derived constants and Prometheus accounting data names used to evaluate the simple payload model for STFC Cloud payloads.

Simple does…

Enhanced should too…

iris Carbon Mapping Project

Scope 2 - Energy	$$E^t_{s\,user} = \frac{S_{user}}{S_{Total}} \cdot E^t_s$$
Scope 3 - Carbon	$$C^t_{e\,s\,user} = \frac{S_{user}}{S_{Total}} \cdot t \cdot Q_{es}$$ Where: $$Q_{es} = \sum_{x=1}^{storage\_items} \frac{C_{ex}}{T_x}$$

Table 5: Summary of the Simple Storage Model showing allocations of Scope 2 energy and Scope 3 carbon to user storage use and the remaining allocation to the provider.

Apportion by quota

Simple Storage Model

Input	Description
E^t_s	Storage Energy usage over an accounting period (including cooling) could be estimated from PDU readings multiplied by PUE
S_{user}	Storage capacity allocated to a user
S_{total}	Total storage capacity of the storage subsystem
t	Duration of accounting period
C_{ex}	Inventory Entry: Embedded carbon of each storage item x
T_x	Inventory Entry: expected lifetime of each storage item x

Table 6: Summary of the inputs needed to evaluate the Simple Storage Model.

iris Carbon Mapping Project

Scope 2 - Energy	$$E^t_{s\,user} = \frac{S_{user}}{S_{total}} \cdot P^{idle}_s \cdot t + \left(E^t_s - P^{idle}_s \cdot t\right) \frac{B_{user}}{\sum_{u=1}^{all\_users} B_u}$$
Scope 3 - Carbon	$$C^t_{e\,s\,user} = \frac{S_{user}}{S_{Total}} \cdot t \cdot Q_{es}$$ Where: $$Q_{es} = \sum_{x=1}^{storage\_items} \frac{C_{ex}}{T_x}$$

Table 7: Summary of the Enhanced Storage Model showing allocations of Scope 2 energy and Scop
storage use and the remaining allocation to the provider.

Apportion idle by quota

Apportion active by use fraction

Know your idle power?

Know your bytes?

Input	Description
E^t_s	Storage Energy usage over an accounting period (including cooling) could be estimated from PDU readings multiplied by PUE
P^{idle}_s	Idle power draw of the storage cluster (including cooling) could be estimated from PDU readings during an idle period multiplied by PUE.
S_{user}	Storage capacity allocated to a user
S_{total}	Total Storage capacity of the storage subsystem
t	Duration of accounting period
B_{user}	Bytes read from, or written to, a users storage area
C_{ex}	Inventory Entry: Embedded carbon of each storage item x
T_x	Inventory Entry: expected lifetime of each storage item x

Table 8: Summary of the inputs needed to evaluate the Enhanced Storage Model.

Enhanced Storage Model

Carbon Mapping Project

Testing the Storage Models

User/group	Quota	kWh
atlas	11500	588.8
dune	1100	56.3
belle	1000	51.2
lhcb	300	15.4
t2k.org	250	12.8
fermilab	200	10.2
other	200	10.2
Unallocated	450	23.0
Total	15000	768.0

Table 12: Results of evaluating the Simple Storage Model on QMUL data for the 24 hour period of 2024-03-27

Input	Value	Description
E_s^t	768 kWh	Storage Energy usage over an accounting period. In this example 5 racks of storage drawing 6.4kW/rack for 24 hours.
S_{user}	-	Storage capacity allocated to a user
S_{Total}	15 PB	Total Storage capacity of the storage subsystem
t	86400 s	Duration of accounting period

Table 11: Measurements, constants and settings used to evaluate the Simple Storage model.

Ran the numbers of simple model on QMUL Batch Farm

Should also work for Cloud

Need to extract per user usage figures for Enhanced model

iris Carbon Mapping Project

Which Payload Model is Best?

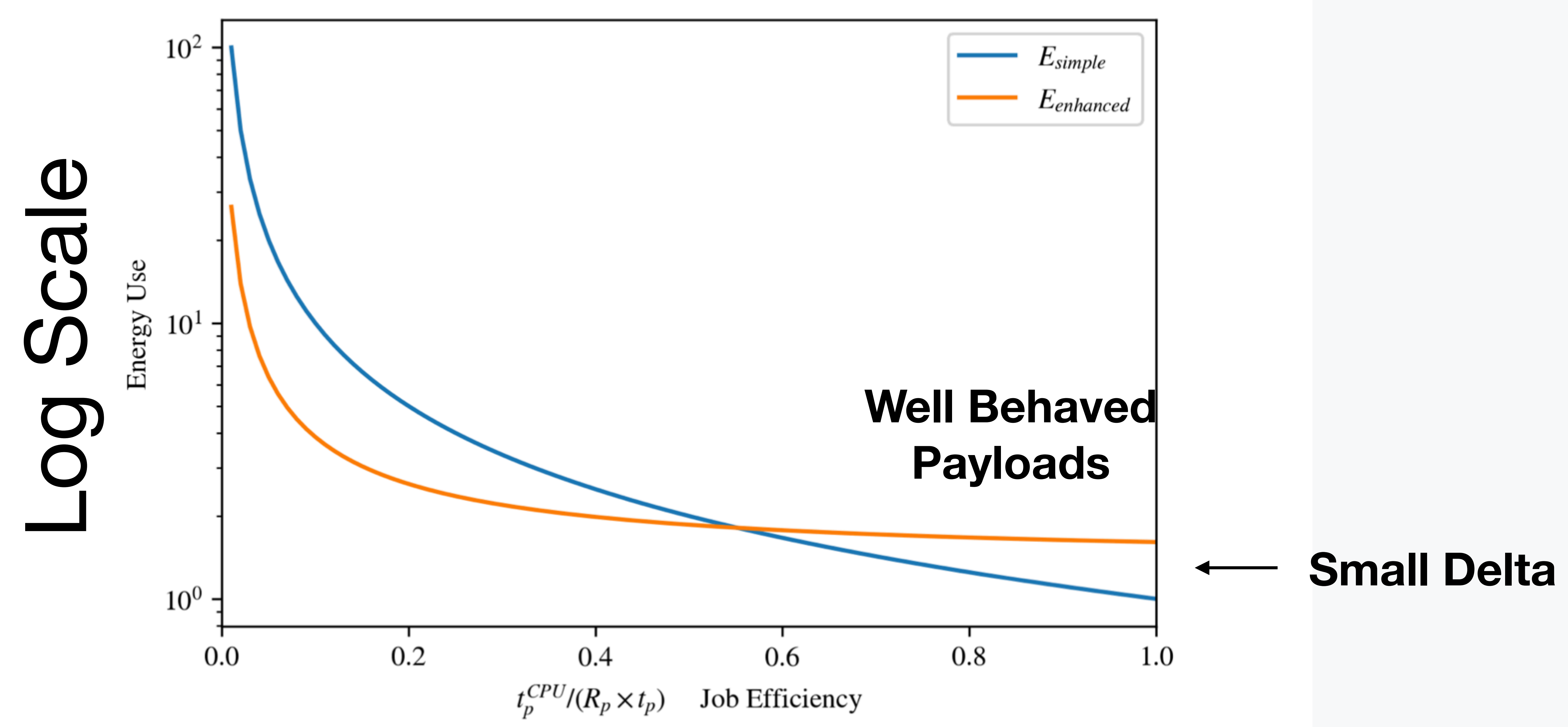


Figure 3: Behaviour of the Simple and Enhanced payload models for a fixed amount of work (constant CPUtime) varying with Job Efficiency. Plotted on a log scale.

Not much to choose between them.

Both encourage more efficient code

Enhanced reduces Allocation to Providers

Which Storage Mode is Best? -> Can we get bytes read/written?

Allocating Carbon Costs to Computing Payloads across Heterogeneous Infrastructures: Final Thoughts…

GPU's Were out of scope

Idle Power is important

Start simple & learn by doing

Measure
Model
Monitor
Moderate

Think like an accountant

Lookout for
NetDRIVE

Thanks for listening

Speaker: Dr Alex Owen

r.a.owen@qmul.ac.uk

Queen Mary University of London

Allocating Carbon Costs to Computing Payloads across Heterogeneous Infrastructures.

Backup Slides

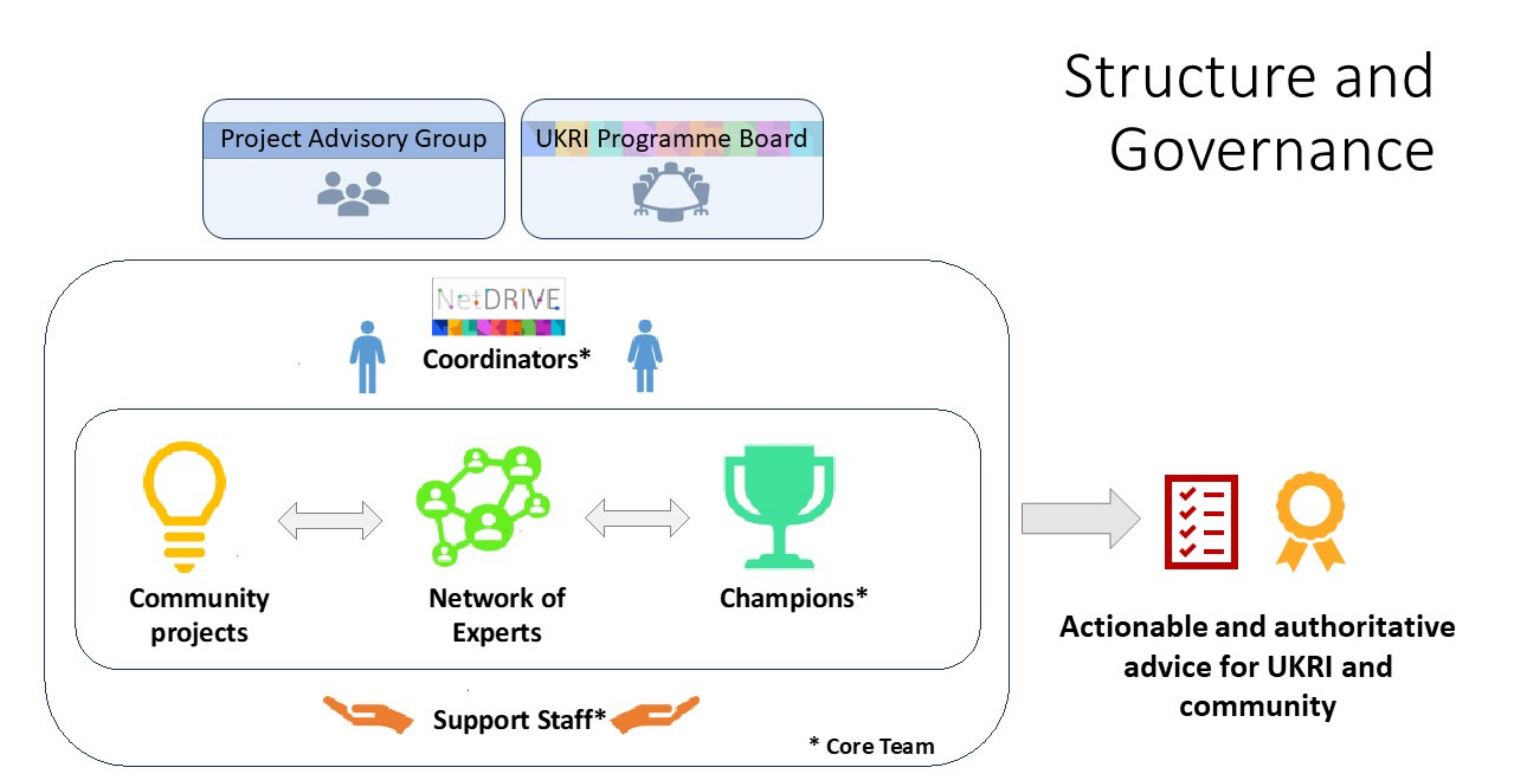
Speaker: Dr Alex Owen

r.a.owen@qmul.ac.uk

Queen Mary University of London

UKRI Delivery Project

Planning now for ~£3M over 3 years starting ~Jan 2025



https://eng.ox.ac.uk/netdrive

iris Carbon Mapping Project

Alex Owen (QMUL)
Jon Hays (QMUL)
Daohai Li (QMUL)
Alex DIbbo (STFC)

Motivation: How should IRIS work towards NetZero DRI?

Allocate Carbon Costs to User Payloads

Reporting Requirements

Outline Delivery Roadmap

Queen Mary University of London
GridPP
UKRI Science and Technology Facilities Council | Scientific Computing

iris Carbon Mapping Project

Reporting Requirements / Concerns

Federation

Carbon costs of IRIS activity/providers broken down into scope 2 and scope 3.

Carbon costs of IRIS supported projects broken down into scope 2 and scope 3.

Carbon saved by being a federation

Reporting upwards:
 Benefit realisation, infrastructure efficiency
 Demonstrate right mix of platforms/tech
 Value of heterogeneity in the federation

Present success while continuing research
 Power used per hepspec
 Fossil power used per hepspc

Try to lead the narrative

Providers

Carbon costs of a provider's service broken down by scope.

Allocate service carbon cost to users and idle/provider

Ease of implementation

Allocate maximum to users (minimum to idle/provider)

Users

Energy per job

Average IRIS Carbon Intensity

Average Embedded carbon factor

Try to avoid motivating behaviour that increase federation carbon costs.

iris Carbon Mapping Project

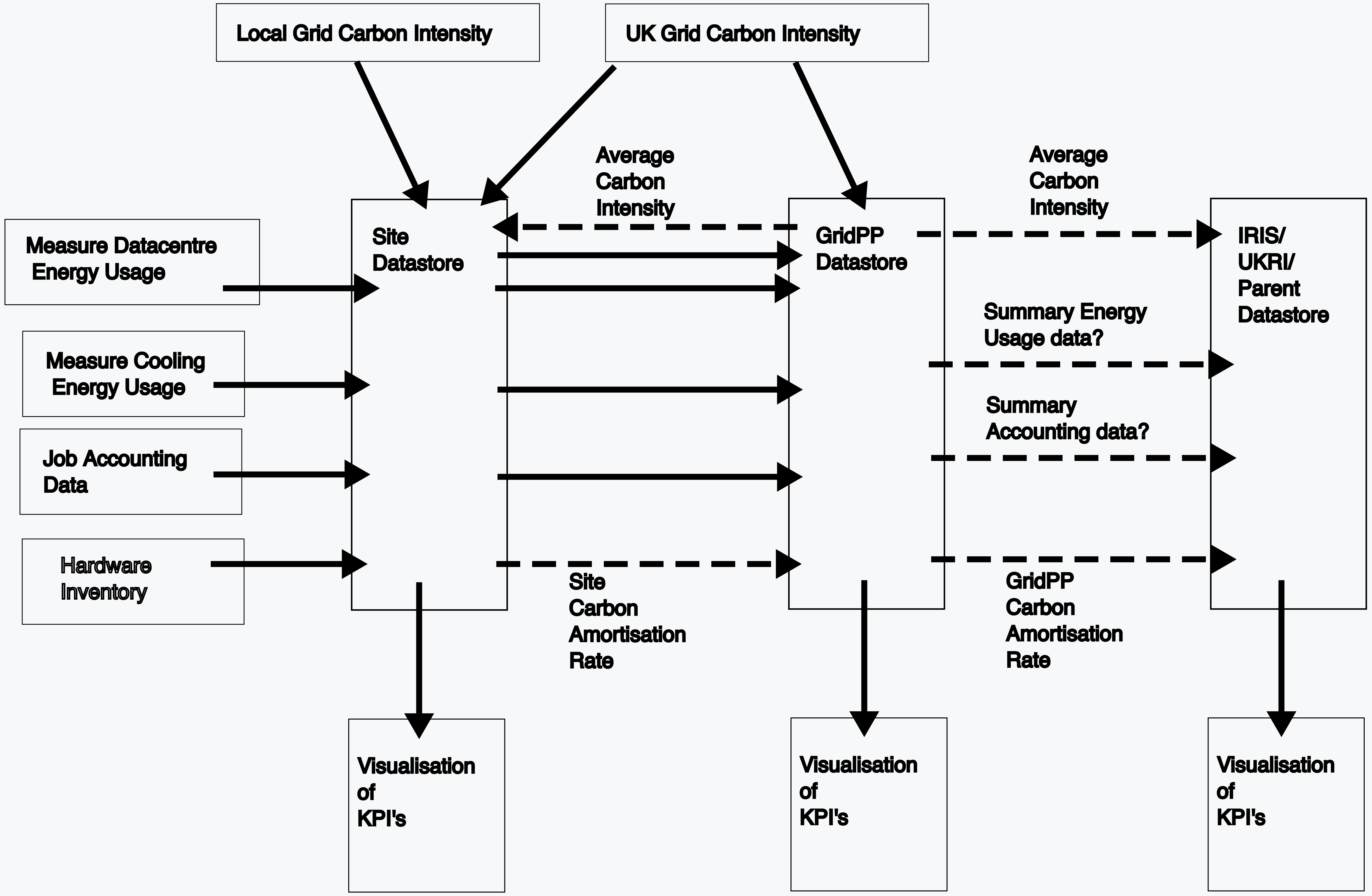
Outline Roadmap

ID	Action	By whom	Timeframe
1	Include energy efficiency and scope 3 carbon considerations into procurements with low weighting	Provider	Now
2	Request LCA and scope 3 data from suppliers at procurement	Provider	Now
3	Increase weighting of energy efficiency and scope 3 carbon considerations into procurements	Provider	Soon
4	Require LCA and scope 3 data from suppliers at procurement	Provider	Later
5	Agree a minimum Carbon Inventory schema	Federation	Now
6	Create and maintain the Carbon Inventory	Provider	Now
7	Decide carbon accounting policy for scope 3 write-off/credit if equipment disposed of early or sold as working	Federation	Now
8	Prepare guidelines on how to optimise lifetime of kit for carbon emissions	Federation	Soon
9	Collect Grid Carbon Intensity for: provider sites, federation average and UK average.	Fed/Prov	Now
10	Publish average federation carbon intensity	Federation	Now
11	Share good practice on how real vs apparent AC power measurements effect the processing of different energy use measurements.	Federation	Now
12	Decide on initial carbon model for payload allocation	Federation	Now

ID	Action	By whom	Timeframe
13	Commission an IRIS Carbon Accounting Data Repository: planning and implementation, including data model and data transfer.	Federation	Now
14	Evaluate selected model on payloads daily to give user energy feedback	Provider	Now
15	Evaluate selected model on payloads monthly to report sum of payload energies and idle energy and apportioned embedded carbon costs	Prov/Fed	Now
16	Collect monthly returns of data from providers to IRIS Carbon Accounting Data Repository	Federation	Now
17	Commission reporting portal to provide the identified reports to federation, providers, and users.	Federation	Now
18	Commission reporting to users of payload energy usage and average federation carbon intensity.	Federation	Now
19	Additional tools for user code optimisation such as energy benchmark tools and the addition of profiling queues to services run by providers.	Fed/Prov	Soon
20	Find or commission an energy benchmark for providers to run on compute nodes and keep results in inventory	Federation	Soon
21	Survey GPU energy monitoring frameworks and plan how to add accelerators into carbon monitoring models.	Federation	Soon
22	Review evidence from under-clocking of accelerators and the effect on carbon emissions.	Federation	Soon
23	Collect additional user carbon reporting needs.	Users	Soon
24	Plan how to record and report the impact of Green RSE's.	Federation	Now
25	Regular review of developments in 'Green Scheduling'.	Federation	Now
26	Regular review of UKRU DRI NetZero projects and policy	Federation	Now
27	Bid for UKRI DRI NetZero funds	ALL	Now
28	Prepare IRIS Carbon Costing Framework for grant proposals	Federation	Now

https://doi.org/10.5281/zenodo.10966001

GridPP ECO-Grid Model



V0.0