Investigating the impact of policy on energy consumption in the Energy Data Centre
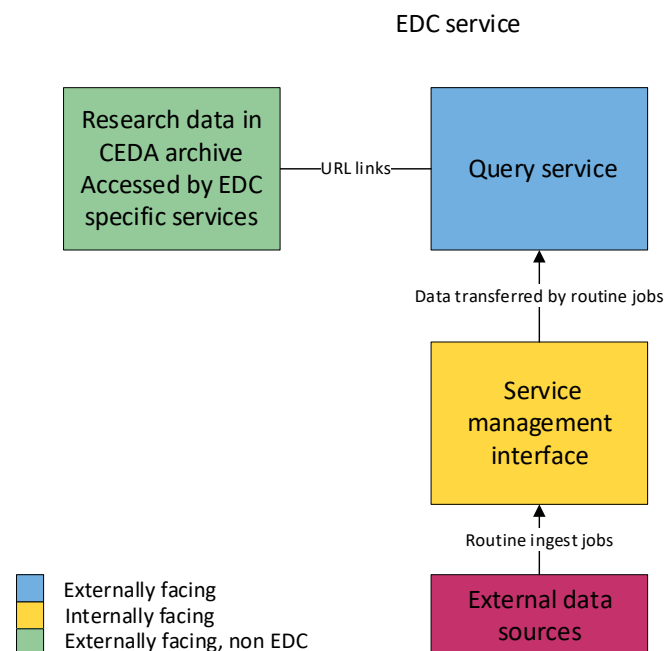
Catherine Jones,
Energy Research Unit, Technology Department,
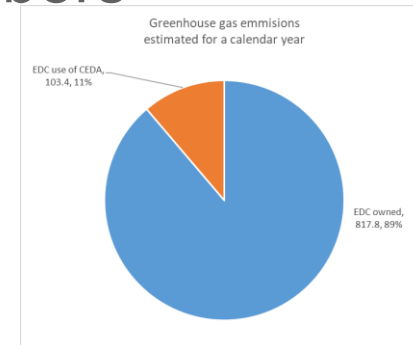UKRI/STFC
DPC Nov 2021
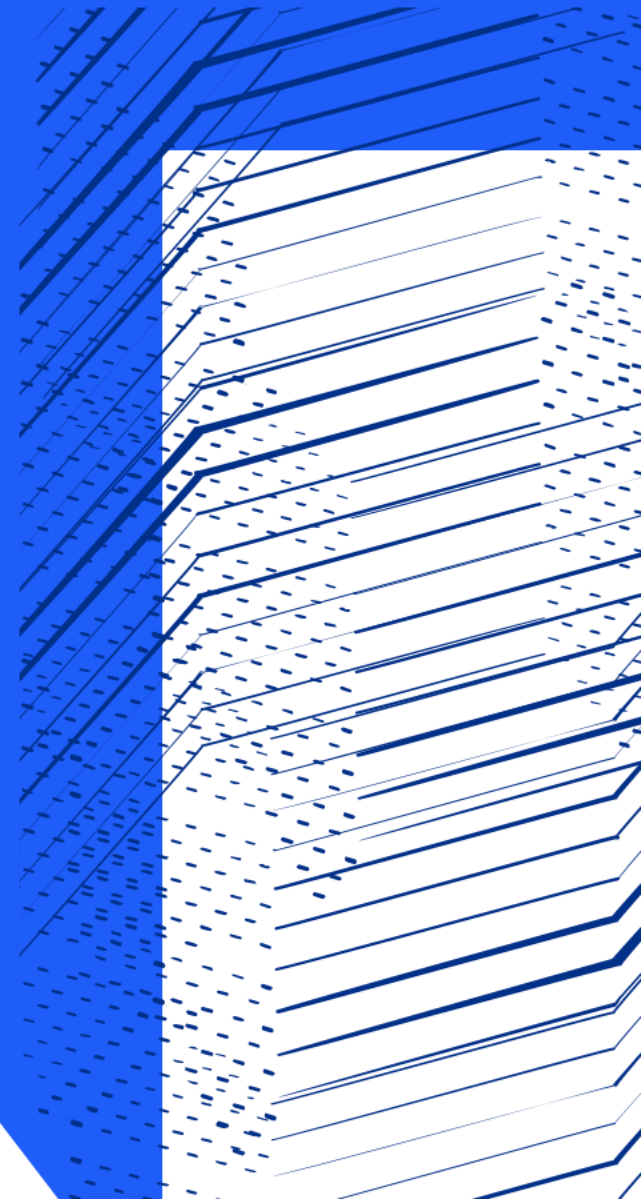
Energy Data Centre: introduction

- Part of the UK Energy Research Centre, funded by UKRI

- UKERC vision: Independent whole systems research for a sustainable energy future

- EDC collects & disseminates: research data; publications and metadata on energy projects

- Infrastructure:
  - Small team: 3 FTEs
  - Internal systems

EDC service

Research data in CEDA archive Accessed by EDC specific services — URL links — Query service

Data transferred by routine jobs

Service management interface

Routine ingest jobs

External data sources

- Externally facing
- Internally facing
- Externally facing, non EDC

UK RI

Science and Technology Facilities Council

Internal Project

- **How does our policy affect our energy usage?**

- **Reviewed service aspirations**
  - DPC RAM & TNA DiAGRAM (risk scenarios)

- **Measured energy usage for routine jobs**
  - Dell servers: IPMI tool (standard for sensor data)
  - Analysed the outputs – did observe changes but very small numbers
  - Worked out greenhouse gas emissions for the service
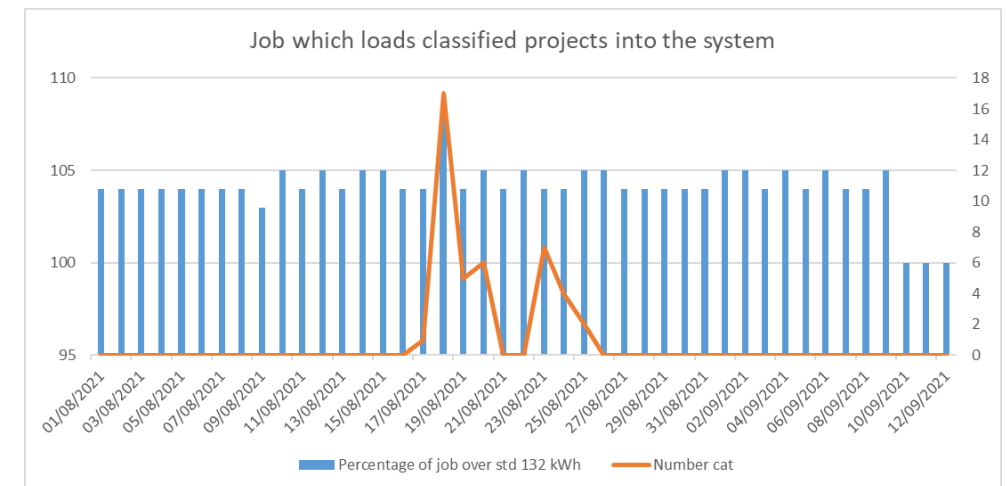 - Also ran new URL checking job

Greenhouse gas emmisions estimated for a calendar year

EDC use of CEDA, 103.4, 11%

EDC owned, 817.8, 89%

All contributions make a small difference and put together makes a bigger difference.
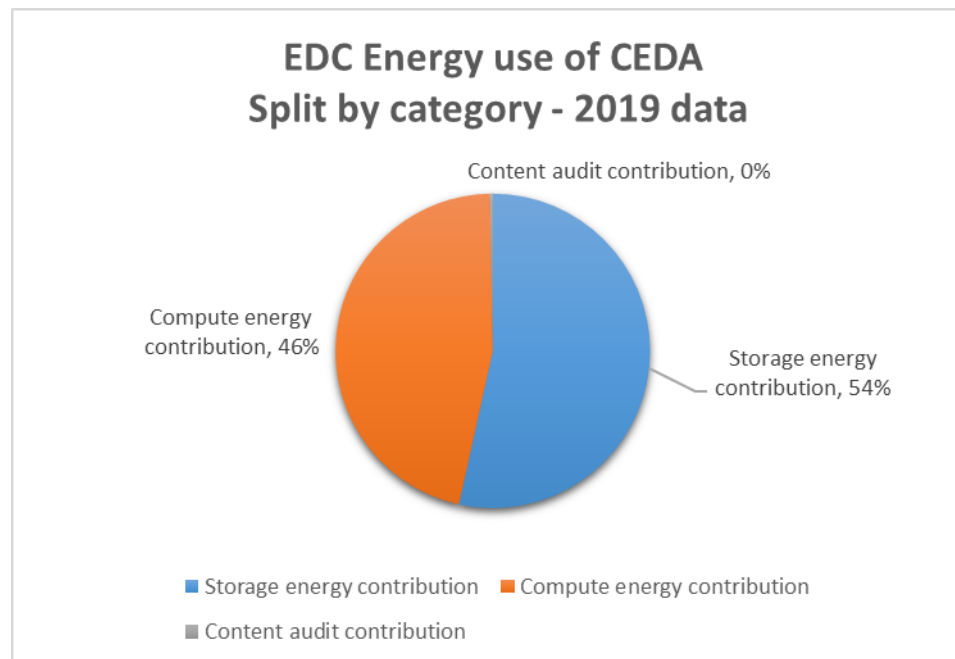
This is equivalent to average car driving **3280** miles a year

Science and Technology Facilities Council

Routine jobs

- Ingest and data quality activities (known times)
- Measured baseline when jobs weren't running.
- Able to measure changes in power consumption – very small numbers
- IPMI provides averages in steps: better data quality on jobs that took longer
- Highlighted some jobs which run automatically but don't always do anything if no new records
- Regular URL checking now in place as a result
- Accidentally left the energy data gathering going so now have 8 months worth of data!

URL checking

Recalculation job
dependant on new data being added

- Percentage of job over std 132 kWh
- No. new added

Job which loads classified projects into the system

- Percentage of job over std 132 kWh
- Number cat
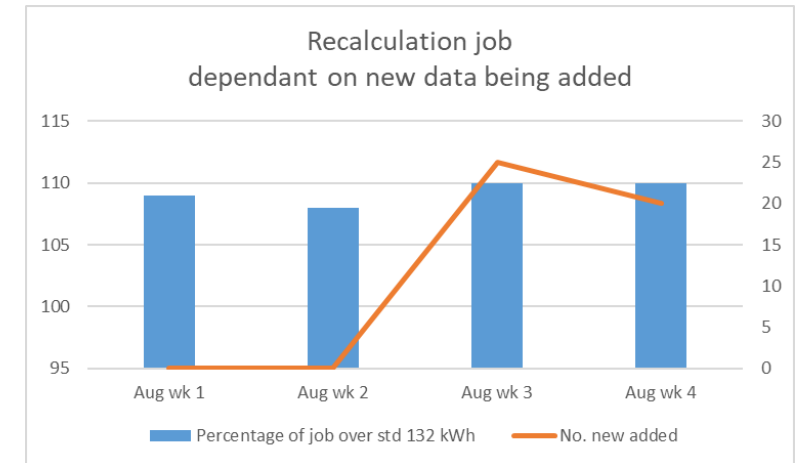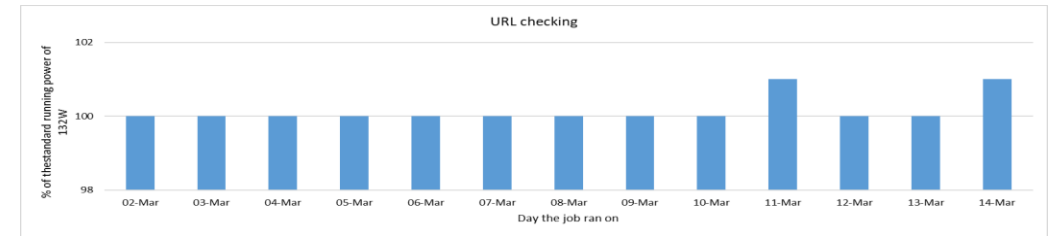
CEDA Component – research data

EDC Energy use of CEDA
Split by category - 2019 data

Content audit contribution, 0%

Compute energy contribution, 46%

Storage energy contribution, 54%

- ■ Storage energy contribution   ■ Compute energy contribution
- ■ Content audit contribution

Not very much energy used in this part of the service – as expected. Would increase as data volumes increase.
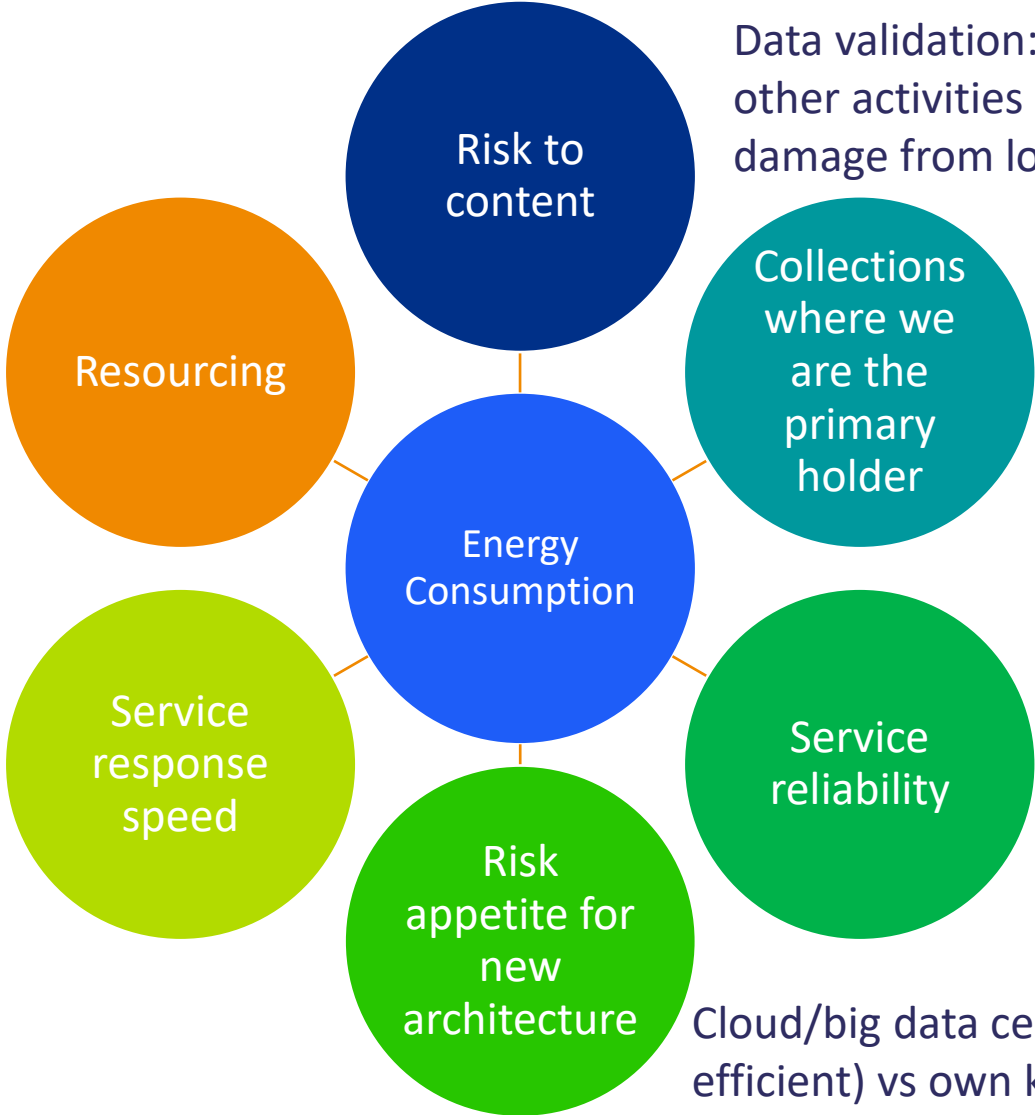Big computing (cloud/virtual) is more energy efficient than a couple of machines for a particular service

Calculated by working out proportion of storage & compute and knowing total energy consumption on JASMIN.
Figures & calculations provided by CEDA staff

Science and Technology Facilities Council

UK RI

Risk benefit discussion

Data validation: checksums & other activities vs reputational damage from losing something

Systems re-write /adaptions to be greener: needs technical effort

Reputational risk (& different processes/risk appetite)

Data on disk (fast) vs on taps (slower) but more energy efficient

Architecture reflecting risk appetite AND service specification/money

Cloud/big data centres (energy efficient) vs own kit (more control).

Risk to content

Resourcing

Collections where we are the primary holder

Energy Consumption

Service response speed

Service reliability

Risk appetite for new architecture

Science and Technology Facilities Council

Conclusions

- Measured energy consumption of routine jobs (with caveats)

- Some new preservation risk mitigations won't increase energy consumption

- Demonstrated using large shared resources is more energy efficient (as expected)

- EDC actions:
  - Reviewed regular jobs
  - Investigate checksums processes on low risk objects
  - Investigate greening the web application
  - Consider risk appetite for the infrastructure
  - Add energy consumption dashboard functionality

https://mobile-efficiency-index.com/en/
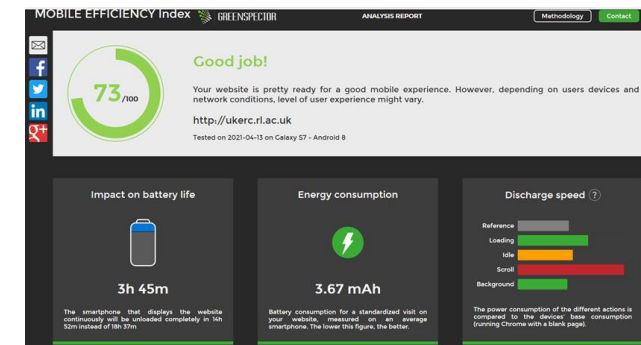
Thanks

- Peter Holt, Sarah James & Alan Ruddell from the Energy Research Unit, Technology Department, UKRI/STFC

- Sam Pepler, Andrew Harwood & Matt Pritchard, Centre for Environmental Data Analytics based at UKRI/STFC

- Chris Dean from UKRI/STFC Digital Infrastructure staff

- Jen Mitcham & Paul Wheatley from DPC for advice & sense checking

- David Underdown from TNA for advice on DiAGRAM

UK RI — Science and Technology Facilities Council

All contributions make a small difference and put together makes a bigger difference.